DOCTOR OF PHILOSOPHY

The Investigation of Oligosaccharyltransferase in Trypanosoma brucei

Jinnelov, Par

Award date:
2014

Division of Biological Chemistry and Drug Discovery
College of Life Science

The Investigation of Oligosaccharyltransferase in *Trypanosoma brucei*

Anders Jinnelöv

Supervisor: Michael A. J. Ferguson

List of figures

List of tables

List of abbreviations

BSA	bovine serum albumin
BSD	blasticidin S deaminase
cKO	conditional knock out
DIG	digoxigenin
dKO	double knock out
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
Dol-PP	dolichol pyrophosphate
DTT	dithiothreeitol
E. coli	Escherichia coli
EDTA	ethylenediamine tetraacetic acid
ER	endoplasmatic reticulum
FSG	fish skin gelatin
Gal	galactose
gDNA	genomic DNA
GDP	guanidine diphosphate
GFP	Green fluorescent protein
Glc	glucose
GlcNAc	N-acetylglucosamine
GMD	GDP-mannose 4,6-dehydratase
GPI	glycosylphosphatidylinositol
GT	glycosyltransferase
HAT	Human African Trypanosomiasis

HYG	hygromycin phosphotransferase
IgG	Immunoglobulin G
IPTG	isopropyl-β-D-thiogalactopyranoside
L. major	*Leishmania major*
LacNAc	N-acetyllactosamine
Man	mannose
mRNA	messanger RNA
ORF	open reading frame
OST	oligosaccharidetransferase
PAC	purimycin acetyltransferase
PCR	polymerase chain reaction
PFA	paraformaldehyde
PMAA	partially methylated alditol acetates
PMSF	phenylmethylsulfonyl fluoride
PNGaseF	peptide:N-glycosidase F
RNA	ribonucleic acid
rRNA	ribosomal RNA
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrilamide gel elecrophoresis
SILAC	Stable isotope labelling with amino acids in cell culture
sKO	single knock-out
Sm	single marker
SSC	sodium chloride, sodium citrate buffer
STT3	Staurosporine and temperature sensitive
T. brucei	*Trypanosoma brucei*
T. cruzi	*Trypanosoma cruzi*

Tet	tetracycline
TEV	Tobacco etch virus
TLCK	tosyl-L-lysine chloromethyl ketone
T_m	melting temperature
Tris	tris(hydroxymethyl)aminomethane
UEA I	*Ulex europaeus* agglutinin I
UTR	untranslated region
VSG	variant surface glycoprotein
YFP	Yellow fluorescent protein

Acknowledgement

Firstly, my greatest thanks go to Mike who, for some reason, gave me the opportunity to come back to his lab and introduced me to a really interesting project and for the supervision, guidance and encouragement throughout my studies.

I also want to give a big thanks to Lucia and Mick who have introduced me to and helped me in many of the techniques used in this thesis. Lucia for cell culturing, immunofluorescence and immunoprecipitations and Mick for general molecular biology.And you both for SILAC guidance. I have learned a lot!

I would like to acknowledge the FingerPrints Proteomics facility and Amy, who helped me with my SILAC IP data and I am also truly grateful for the help and support from the rest of the lab members, both former and present. You have all contributed to a very friendly lab environment. The same for all amazing lab managers.

A special thanks go to my fellow end bay friends and colleague. I honestly do not think it's allowed to laugh as much as we have done during the years. Seb1, vielen dank for alles deine support und suggestions und for alles germlish conversations. Ich hoffe meine Frikadelle hast nicht zu schlecht nase geschmeckt! Angela thanks for all help, spanning from personal taxi when I moved flats to reading what will follow this spasmodic page of writing. Kycklingsoppa? Was that my fault? Anna, thanks for being a true colleague. The non-friend type. Eternally (not really, but it sounds good) thankful for you being my personal calendar, bank and IT support for the times Windows' Help and Support was not enough. Also thanks for being part of Team Eat. Chocolate time?

I'm also grateful to people in other labs, both for helpful scientific discussions and for plain shenanigans. Adam, Han, Susan, Seb-lost Lucy, Dun Jack and Kat, you are all included.

Finally, my love and gratitude to my family and friends in Sweden (and elsewhere) for your unwavering support and encouragement. You, together with the people aforementioned, have played a major part in making this happen (Jinnelov et al., 2014). Thank you all.

Declaration

I declare that I am the author of this thesis; all references cited have been consulted by myself: the work of which this thesis is a record, unless specifically stated, has been done by myself and this work has not been previously accepted for a higher degree.

Anders Jinnelov

I confirm that Anders Jinnelov has performed the research described in this thesis under my supervision and has fulfilled the conditions of the relevant ordinance and regulations of the University of Dundee and that he is qualified to submit the following thesis for the degree of Doctor of Philosophy.

Professor Michael A. J. Ferguson

Abstract

The parasite *Trypanosoma brucei* is the causative agent of Human African trypanosomiasis and Nagana in cattle and millions of people in sub-Saharan Africa are at risk for both health and economic issues since current drugs are inadequate.

N-linked glycosylation in *Trypanosoma brucei*differs from other eukaryotes and the parasite has several essential glycoproteins.Thus, the study of oligosaccharyltransferase activities is interesting for potential future drug discoveries. To assess whether *Tb*STT3A is essential *in vitro*(i.e. in culture), it was replaced by a drug resistance gene in the presence of a tetracycline inducible *Tb*STT3A ectopic copy. Although both endogenous *Tb*STT3A genes were replaced in the *Tb*STT3 locus, the *Tb*STT3A function was shown to be retained. These results might suggest that the cells have rearranged their genome during *Tb*STT3A replacement, which would imply that the tetracycline-inducible ectopic copy could not complement the loss of the endogenous *Tb*STT3A gene, suggesting that it is essential for cell viability *in vitro*. Additionally, *in situ* tagging of *Tb*STT3A enabled us to search for potential binding partners by performing blue native gel electrophoresis and immunoprecipitation using a Stable Isotope Labelling of Amino acids in Cell culture (SILAC) methodology. The results, further strengthened by co-immunoprecipitation, suggested that *Tb*STT3A forms a large multimeric complex with *Tb*STT3B in *Trypanosoma brucei*. To our knowledge, this is the first time two different STT3 enzymes have been reported to associate in an OST complex. Furthermore, to investigate why *Tb*STT3A prefer sequons surrounded by acidic amino acids, protein models of the *Trypanosoma brucei* STT3s were constructed. They revealed an intriguing arginine residue at position 397 (that was not found in*Tb*STT3B), which seems to point straight into the active site of *Tb*STT3A. To further

elucidate how surrounding acidity affects *Tb*STT3A glycan transfer, different acceptor peptides containing only alanines and a single aspartic acid in various flanking positions were fused to a glycosylation reporter protein. Using this reporter, the percentage of *Tb*STT3A glycan transfer was calculated. The results showed that a single aspartic acid in any of the two positions immediately adjacent to the asparagine (X-N-X-S/T) increased *Tb*STT3A glycosylation up to 60-70 % compared to 5 % in the alanine control. Taken together, it is possible that the arginine in position 397 can interact with the aspartic acids immediately adjacent to the asparagine and thus play a major role in *Tb*STT3A and *Tb*STT3B sequon specificity.

The findings from these experiments have given a greater understanding of the *Tb*OST complex and the specific acceptor peptide selectivity of *Tb*STT3A and *Tb*STT3B in *Trypanosoma brucei.*

1 – Introduction

1.1 Human African Trypanosomiasis

Human African trypanosomiasis (HAT), or sleeping sickness, is a vector-borne disease caused by the extracellular protozoan parasite *Trypanosoma brucei* (*T. brucei*). *T. brucei* is transmitted by the bite of a tsetse fly (*Glossina* genus) when taking a blood meal and the resulting disease in cattle is called nagana. The disease is only spread in sub-Saharan Africa, in discrete areas of endemicity within the geographic distribution of the tsetse fly. Humans are infected by two parasite sub-speices: *Trypanosoma brucei rhodesiense* and *Trypanosoma brucei gambiense*(Simarro et al., 2012). *T. b. gambiense* infections are mainly seen in West and Central Africa and cause the chronic form of the disease (Balmer et al., 2011; Checchi et al., 2008a, 2008b). *T. b. rhodesiense* infections, on the other hand, are mainly seen in Eastern and Southern Africa and cause an acute form of the disease with a more rapid onset and progression. (Balmer et al., 2011; Chappuis et al., 2005a; Simarro et al., 2012).*Trypanosoma brucei gambiense* is responsible for 90 to 95 % of all reported cases, with *T. b. rhodesiense* accountable for the remainder (Simarro et al., 2010).Both parasites cause early stage symptoms including fever, malaise, anaemia, headache, weight loss and weakness. In this first stage, the parasites are limited to the bloodstreamand the lymphatic system (Sternberg, 2004). In the later stage, the trypanosomes have crossed the blood-brain barrier and invaded the central nervous system. The latter, more advanced stage, leads to neurological damage, sleep cycle disruption and coma. Both forms of HAT are generally lethal if left untreated (MacLean et al., 2012; Wolburg et al., 2012).

Today, the recorded number of new cases has reduced below 10,000 per year,however the prevalence of HAT is predicted to be higher due to under-reporting. Furthermore, approximately 70 million people are estimated to be at risk of contracting HAT(Simarro

et al., 2011, 2012).The standard field diagnostic tool for HAT is the card agglutination trypanosomiasis test (CATT). However, this test has limitations including limited sensitivity and specificity and only being *T. b. gambiense* compatible. Furthermore. it is also difficult to distinguishing between infected and cured individuals (Lejon et al., 2010; Paquet et al., 1992). Determining the stage of the infection is still necessary to predict whether the trypanosomes have passed the blood-brain barrier and invaded the CNS or not. This is done by lumbar puncture and cerebrospinal fluid examination (Chappuis et al., 2005b).

Many of the drugs that are currently used to treat the disease are old and have severe limitations, whilst none of the therapeutic agents are effective against both stages. pentamidine and suramin can be used against the first stage of HAT caused by *T. b. gambiense*, while first stage HAT caused by *T. b. rhodesiense* can only be treated by suramin (Bouteille et al., 2003).However, these compounds do not penetrate the blood-brain barrier. The drug of choice forlate stage *T. b. gambiense* is mainly nifurtimox-eflornithine combination therapy (NECT), whereas melasoprol is only used for late stage *T. b. rhodesiense* infections (Kennedy, 2012; Priotto et al., 2009; Simarro et al., 2011).Some of the aforementioned therapeutic agentscause severeside-effects, including melasoprolwhich can cause lethal encephalopathy in 5-10% of the cases. (Fairlamb, 2003; Renslo and McKerrow, 2006).The drugs used for treating HAT are also expensive and difficult to administer in rural areas because of the need for sterile instruments and trained personal (Simarro et al., 2008). Therefore, new, affordable and safer drugs that are effective against both stages of infections and both sub-species, and that are easier to administrate, are urgently needed.

1.2 *Trypanosoma brucei* life cycle and cell biology

T. brucei is a unicellular parasite that belongs to the kinetoplastida order, family trypanosomatidae. The parasite has a complex life cycle that allows it to alternate between the insect vector and the mammalian host (Figure 1.1). This requires the parasite to present different morphologies and surface coatsto survive in the different environments (Fenn and Matthews, 2007). Bloodstream form parasites can be ingested by the tsetse fly during a blood meal, and within the fly, the parasites differentiate into the proliferative procyclic form to be able to colonize the insect mid gut. Subsequently, the procyclic parasites migrate to the insect salivary gland, where they differentiate into the epimastigote form. The proliferative epimastigotesthen differentiate into the infective metacyclic trypomastigotes, which are ready to be injected into another mammalian host during a new blood meal (Van Den Abbeele et al., 1999; Sbicego et al., 1999; Sharma et al., 2009). When this occurs, the metacyclic parasites are injected into the bloodstream of the mammalian host, where they differentiate into long and slender bloodstream form trypomastigotes, dividing rapidly and infecting other sites. However, the bloodstream cell population is phleomorphic, and when the cell concentration increases more cells develop into the stumpy form. The stumpy form is non-proliferative and ready to be taken up by the insect via a new blood meal and subsequently differentiate to procyclic form, thereby continuing the life cycle (Fenn and Matthews, 2007; Sbicego et al., 1999).

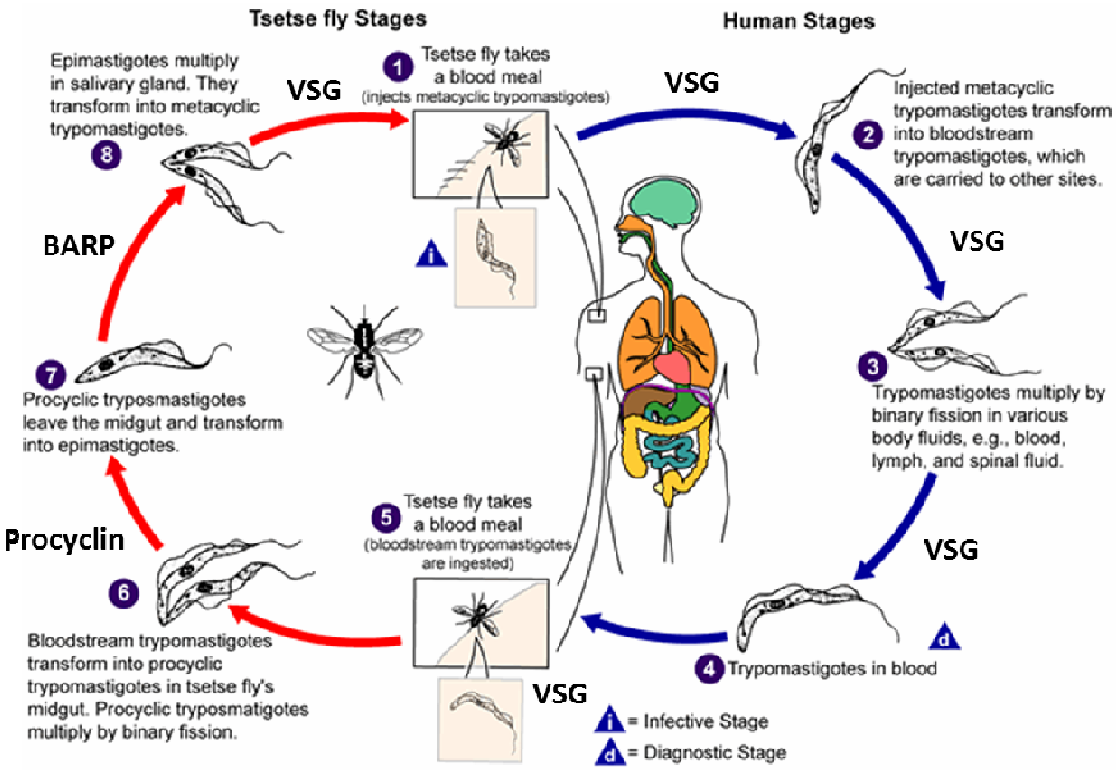


Figure 1.1. Life cycle of *T. brucei*. Image adapted from Centre for Disease Control (http://www.cdc.gov/parasites/sleepingsickness/biology.html)

The bloodstream form of the parasite is covered by a dense coat of variant surface glycoproteins (VSG) which allows small molecules to diffuse to the plasma membrane, but prevents macromolecules from reaching it, e.g. the parasites are protected from complement mediated lysis.(Cross, 1975a; Fenn and Matthews, 2007). The surface coat changes when stumpy bloodstream parasites differentiate into the procyclic form, which has procyclin glycoproteins as the most abundant surface molecule (Sbicego et al., 1999). The epimastigote form, on the other hand, expresses a surface coat protein called brucei alanine rich protein (BARP) (Fenn and Matthews, 2007).

The experiments in this thesis have solely been done using the bloodstream form *T. brucei* and therefore the cell biology will be focussed on this specific stage of the life cycle.

The cell surface of bloodstream form *Trypanosoma brucei* is covered with an estimated 5 million homodimers of VSG glycoproteins. The parasite has several hundred genes, some of which are pseudogenes, coding for these VSG dimers, allowing the parasite to switch between immunologically distinct VSG surface coats. This antigenic variation is the reason why the host immune system cannot control parasite growth and why the parasitemia occurs in waves (Cross, 1975a, 1979; Ferguson et al., 1988; Horn and Cross, 1997).

T. brucei is a flagellated organism, where the main function of the flagellum is motility, which has been shown to be essential for the parasite viability(Broadhead et al., 2006). The flagellum exits the cell in the flagella pocket, which is the sole known site for endocytosis and exocytosis even though it only represents approximately 5 % of the plasma membrane (Landfear and Ignatushchenko, 2001).

T. brucei parasite cells have an organelle called the glycosome, which isthe trypanosome equivalent of the peroxisome. However, unlike other eukaryotes, most of the steps of glycolysis occur inside this organelle (Michels et al., 2006; Opperdoes and Borst, 1977). Several biosynthetic pathways have been localised in the glycosome, including, β-oxidation of fatty acids (which is commonly found in peroxomes) the pentose-phosphate pathway, purine salvage and pyrimidine biosynthesis (Cáceres et al., 2010; Michels et al., 2006; Zomer et al., 1995).What is particularly relevant to work in this thesis, is that most, if not all, of the sugar nucleotide biosynthesis is localised to the glycosome. These activated sugar nucleotides play an important role in N-glycan assembly (Bandini et al., 2012; Kuettel et al., 2012; Mariño et al., 2010; Roper et al., 2005; Stokes et al., 2008; Turnock and Ferguson, 2007).

1.3 N-linked glycobiology

Asparagine (N)-linked glycosylation of proteins is a vital and extensive post-translational modification catalysed by the oligosaccharyltransferase (OST). Glycosyaltion results in a covalent bond between a largepreassembled glycan and the asparagine residue in a specific sequon, N-X-S/T (variations of the sequon occurs, see section 1.3.2), found in nascent proteins. For membrane and secretory proteins in eukaryotic organisms, it is the mostfrequent protein modification. N-linked glycans can serve numerous functions including aiding the folding process of proteins in the ER (Helenius, 1994). The hydrophilic nature of the covalently attached glycan alters the biophysical properties of the nascent protein and can affect the conformation of the unfolded polypeptide chain and it can thereby influence protein folding (Wormald and Dwek, 1999). Furthermore, glycans can mediate interactions of the cell or the organism with its environment (Sharon and Lis, 1995).

The high conservation of the process across all domains of life has made it feasible to identify the major principles of the pathway of N-linked glycosylation (Lizak et al., 2011; Wacker et al., 2006).

1.3.1 Glycan assembly in eukaryotes

In most eukaryotes, the oligosaccharide assembly starts on an isoprenoid lipid at the cytoplasmic side of the ER membrane. The isoprenoid lipid, the dolichol molecule, carries the oligosaccharide in the membrane of the ER, making it a vital part of the N-linked glycosylation pathway (Lehrman, 1991; Schenk et al., 2001). The dolichol molecule is synthesised from farnesylpyrophosphate by the sequential addition of C5

isoprenoid units, catalysed by the *cis*-prenyltransferase. The length of the dolichol carrier is species-dependent and the trypanosamatid version is unusually short, containing 11-12isoprene units (Figure 1.2) (Low et al., 1991; Swiezewska and Danikiewicz, 2005; Welti, 2013).

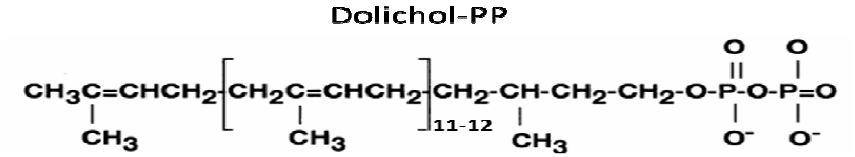


Figure 1.2. The dolichol molecule found involved in oligosaccharide assembly in *T. brucei*.

Dolichol is subsequently phosphorylated by the Dol kinase (DOLK) and is then ready to serve as a carrier for the precursor oligosaccharide assembly (Heller et al., 1992). The start of oligosaccharide precursor assembly is located to the cytoplasmic side of the ER membrane and three different nucleotide-activated sugar donors are needed to synthesise the $Man_5GlcNAc_2$ intermediate oligosaccharide; UDP-*N*-acetylglucosamine (UDP-GlcNAc), GDP-mannose (GDP-Man) and UDP-glucose (UDP-Glc) (Burda and Aebi, 1999; Huffaker and Robbins, 1983).These sugar nucleotides act as substrates for glycosylation reactions, andin higher eukaryotesthey are synthesised in the cytoplasm and subsequently used on the ER surface and in the lumen, as well as the Golgi apparatus. (Varki A, Esko JD, 2009).A number of glycosyltransferases are involved in these sugar donor transfers and ALG7 and ALG13/14 catalyse the transferof the first two GlcNAc sugars from UDP-GlcNAc to the Dol-P molecule. In the first of these two reactions, GlcNAc is transferred together with a phosphategroup, generating dolichol-PP-GlcNAc. Following the addition of the second GlcNAc, five Man sugars are transferred to the growing lipid-linked oligosaccharide (LLO). These reactions are

catalysed by ALG1, ALG2 and ALG11, and use GDP-Man as their substrate(Cipollo et al., 2001; Couto et al., 1984; O'Reilly et al., 2006). The $Man_5GlcNAc_2$-PP-Dol molecule is then flipped to face the lumenal side of the ER. Some evidence suggested that the Rft1 protein in yeast was involved in translocation of the LLO intermediate, however, a subsequent *in vitro* analysis challenged these results (Frank et al., 2008; Helenius and Aebi, 2002; Ng et al., 2000; Rush et al., 2009).Work carried out in the *T. brucei* procyclic form has suggested that Rft1 influences assembly of the LLO precursor, although not by catalysing the translocation through the ER membrane (Jelk et al., 2013). On the lumenal side of the ER, assembly continues by the addition of four mannose sugars to form the B and C branches (Figure 1.3). These reactions are catalysed by ALG3, ALG9 and ALG12(Aebi et al., 1996; Burda et al., 1999; Frank and Aebi, 2005). In these steps, Dol-P also plays a crucial role as a carrier for Man and Glc which are transferred from GDP-Man and UDP-Glc to the Dol-P molecule in order to transfer these activated sugar nucleotides to the assembling precursor glycan on the lumenal side of the ER membrane (Heesen et al., 1994; Orlean et al., 1988). How the dolichol-Pis able to flip across the membrane with bound Man or Glc is yet to be discovered.In most eukaryotes, the final step of lipid-linked oligosaccharide assembly is the addition of three glucose molecules to the A branch from Dolichol-P-Glc by AlG6, ALG8 and ALG10 (Burda and Aebi, 1998; Reiss et al., 1996; Stagljar et al., 1994). Thus, the precursor oligosaccharide in most eukaryotes is $Glc_3Man_9GlcNAc_2$ (Figure 1.4).

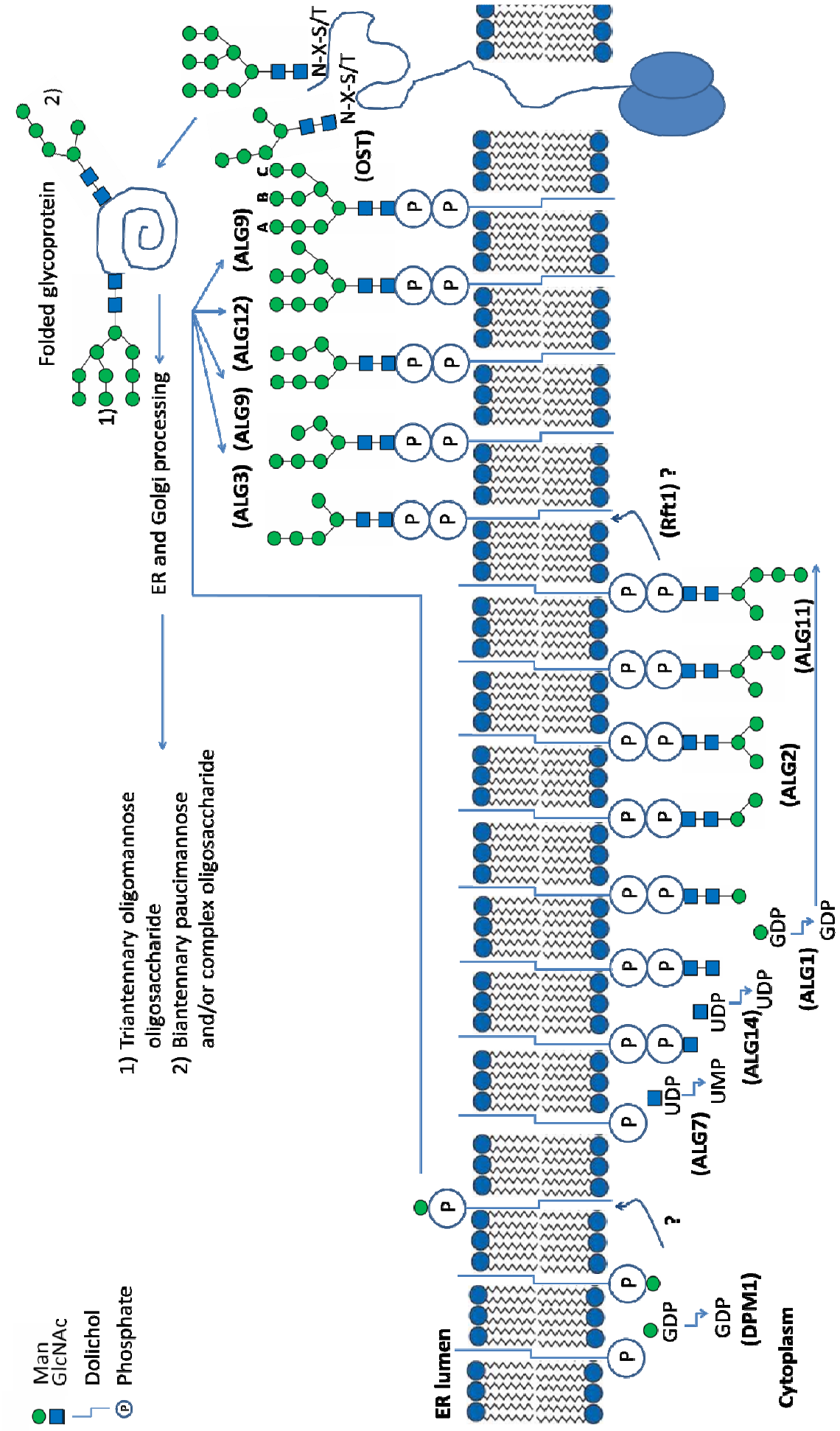

Man
GlcNAc
Dolichol
Phosphate
ER lumen
Cytoplasm
GDP
(DPM1)
?
(ALG7)
UDP
UMP
(ALG14)
UDP
(ALG1)
GDP
(ALG2)
(ALG11)
(Rft1) ?
(ALG3)
(ALG9)
(ALG12)
(ALG9)
A
B
C
(OST)
N-X-S/T
N-X-S/T
Folded glycoprotein
1)
2)
ER and Golgi processing
1) Triantennary oligomannose oligosaccharide
2) Biantennary paucimannose and/or complex oligosaccharide

Figure 1.3

Figure 1.3.(previous page) Schematic view of lipid-linked oligosaccharide assembly in *T. brucei.* The first GlcNAc is transferred as GlcNAc-P from UDP-GlcNAc in a reaction catalysed by ALG7. In the second reaction, GlcNAc is transferred from UPD-GlcNAc to the growing LLO. The remaining five reactions on the cytoplasmic side of the ER are the additions of Man residues from GPD-Man in reactions catalysed by ALG1, ALG2 and ALG11. The $Man_5GlcNAc_2$-PP-Dol molecule is then flipped to the lumenal side of the ER, in a reaction that is suggested to involve Rft1. On the lumenal side, $Man_5GlcNAc_2$-PP-Dol is extended to $Man_9GlcNAc_2$-PP-Dol by the addition of four mannoses from Man-P-Dol to the B and C branches (different branches highlighted for $Man_9GlcNAc_2$), catalysed by ALG3, ALG9 and ALG12. The Dol-P-Man donor is made by the transfer of Man from GDP-Man to Dol-P by DPMS1 on the cytoplasmic face of the ER and is translocated into the lumen of the ER. Subsequently, OSTs transfer $Man_5GlcNAc_2$ and $Man_9GlcNAc_2$ to N-X-S/T sequons in nascent proteins, which aids protein folding. The oligosaccharides can be processed into different types of glycans in the ER and Golgi. In most eukaryotes, the final step of lipid-linked oligosaccharide assembly is the addition of three glucose molecules to the A branch from Dolichol-P-Glc by ALG6, ALG8 and ALG10.

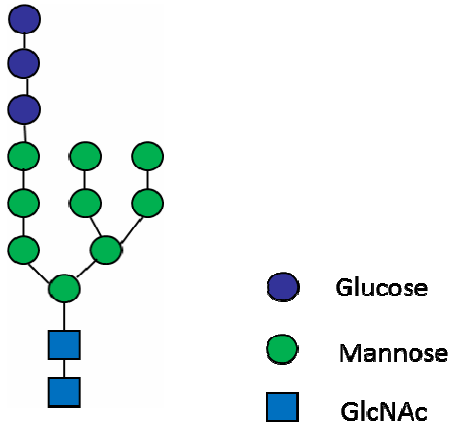


Figure 1.4. Precursor oligosaccharide for most eukaryotes, $Glc_3Man_9GlcNAc_2$.

While most of the ALG genes have identifiable homologues in *T. brucei*, only two have been experimentally characterised; *Tb*ALG3 and *Tb*ALG12 (Jones et al., 2005; Manthri et al., 2008; Samuelson et al., 2005). Notably, dolichol-P-Glc dependent glucosylation reactions do not occur in the parasite and $Man_9GlcNAc_2$-PP-Dol is the largest precursor oligosaccharide made by *T. brucei* (Figure 1.3) (Acosta-Serrano et al., 2004; Jones et al., 2005).

The dolichol carrier can be regenerated following glycosylation which means that the levels of the dolichol molecule on the cytoplasmic side of the ER are not solely dependent on the *de novo* biosynthesis but also the recycling of the molecule (van Berkel et al., 1999).

1.3.2 Acceptor peptide sequon

When nascent proteins enter the ER lumen, they are in an extended and flexible conformation which is preferred by OST for efficient transfer of the precursor oligosaccharides. Subsequently, N-glycosylation occurs approximately 30 residues from where the ribosomes meets the translocon on the ER surface (Varki et al., 1999). A later study refined the knowledge to show that glycosylation occur approximately 65 residues from the peptidyltransferase centre in the ribosome (Kowarik et al., 2002).However, N-glycans arenot transferred to everypotential glycosylation sequon and some sites are sub-stoichiometrically occupied (Jenkins et al., 1996). The consensus of the field has been that up to approximately 60 % of all proteins are N-glycosylated, but recent data challenge that view, stating that less than one fifth of all proteins actually receive glycans from OSTs (Apweiler et al., 1999; Khoury et al., 2011). Furthermore, it has been estimated that approximately 70 % of the sequons on mouse glycoproteins are occupied (Zielinska et al., 2010).

Even though the N-X-S/T sequon is the most abundant in N-glycan transfer, there are exceptions. Bacterial OST transfers lipid-linked oligosaccharides to D/E-X-N-X-S/T sequons, thus exhibiting a more stringent specificity for acceptor sequons than seen in eukaryotes. Additionally, 1 % of the glycosylation sequons that are efficiently used in the mouse glycoproteome are N-X-C sites (Chen et al., 2007; Wacker et al., 2006; Zielinska et al., 2012). Also, valine, leucine or asparagine can switch with the serine in the sequon of a cell surface protein in *Halobacterium halobium* and an N-S-G sequon has been shown to accept an oligosaccharide in *Campylobacter lari*(Schwarz et al., 2011; Zeitler et al., 1998). Furthermore, detailed analysis has suggested that threonine is found more often than serine in the +2 position of the N-X-S/T sequon and that small hydrophobic amino acids are preferred at the X position (Zielinska et al., 2012). Lastly,

it has been proposed that N-glycosylation sequons are mostly found in potential loops and turns of proteins (Petrescu et al., 2004).

1.3.3 Oligosaccharyltransferase – structure and function

Ina crucial reaction of the N-linked glycosylation pathway, OST, a multimeric complex in mosteukaryotes, transfers a preassembled precursor oligosaccharide to the asparagine residues within the sequons of nascent proteins. Since OSTs glycosylate polypeptides that enter the ER, it was not surprising that data provided evidence that parts of the OST are associated with the translocon (Shibatani et al., 2005).

Human OST is a heptameric membrane complex (Figure 1.5) whereas yeast OST, the most extensively investigated, consists of eight subunits: Ost1 (ribophorin I), Ost2 (DAD1), Ost3 or Ost6 (N33/Tusc3 or IAP/MagT1), Ost4 (Ost4), Ost5, Wbp1 (OST48), Swp1 (ribophorin II) and STT3 (STT3A/STT3B) (human subunits in brackets). Of these subunit proteins, fiveare essential for yeast viability: Ost1 (ribophorin I), Ost2 (DAD1), Wbp1 (OST48), Swp1 (ribophorin II) and STT3 (STT3A/STT3B) (Kelleher and Gilmore, 2006; Li et al., 2008; Mohorko et al., 2011).

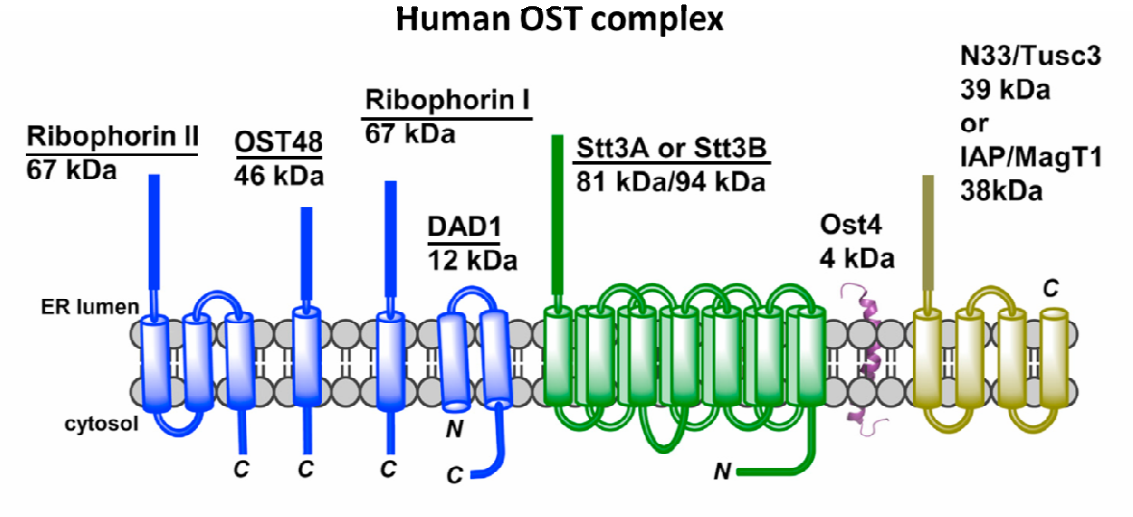


Figure 1.5. Subunits of human OST (Mohoker et al, 2014).

Ribophorin I, ribophorin II and OST48 were the first protein subunits of the mammalian OST complex to be identified (Silberstein et al., 1992). Ribophorin I has one predicted transmembrane helix; a small C-terminal cytosolic domain and a large, luminal domain. Itis an abundant rough ER proteinthat is approximatelystoichiometric to membrane-bound ribosomes (Kreibich et al., 1978a, 1978b; Marcantonio et al., 1984). Additionally, ribophorin I has been shown to enhancethe N-glycosylation of specific type I and type II membrane proteins, possibly by binding poorly glycosylated sequons, thereby giving these peptides a greater chanceto be acted upon by the catalytic STT3 subunits of the OST complex. (Wilson and High, 2007).

Ribophorin II is also found in approximately stoichiometric amounts with ribophorin I and membrane-bound ribosomes(Kreibich et al., 1978a, 1978b; Marcantonio et al., 1984). It is highly conserved in mammals and sequence analysis suggests it has a large N-terminal lumenal domain, three predicted α-helices in a transmembrane domain and a small C-terminal cytosolic segment (Crimaudo et al., 1987). Downregulation of ribophorin II in human breast cancer cells has been shown to lead tohypoglycosylation of selected membrane proteins(Honma et al., 2008).

Ost48 was detected together with the two ribophorins (Silberstein et al., 1992). Human Ost48 has a large N-terminal lumenal domain, followed by a single transmembrane span and a short C-terminal cytoplasmic domain consisting of nine amino acids. The C-terminus of canine OST48 has been shown to interact with DAD1, whereas the N-terminal lumenal domain has been reported to interact with ribophorin I in yeast. In fact, Ost48 seems to act as a link between ribophorin I and ribophorin II by interacting with the lumenal domains of the ribophorins, which have not been reported to interact directly (Fu et al., 1997). Additionally, Wbp1 in yeast has been suggested to bind the LLO substrate, but convincing experimental evidence of this notion has not been presented (Pathak et al., 1995).

DAD1 is the most conserved OST subunit and the human protein is highly similar to the mouse, rat and hamster homologues. The protein (12 kDa) is the second smallest subunit of the OST complex, with both termini on the cytosolic side of the ER membrane and two transmembrane helices in the centre. As mentioned above, the DAD1 C-terminal domain has been reported to interact with OST48 and, moreover, the protein has been reported to be in tight association with the active OST complex (Fu et al., 1997; Kelleher and Gilmore, 1997).

Ost4 is the smallest protein subunit in both mammalian and yeast OST complex, consisting of only 37 residues including a single transmembrane helix (Kelleher and Gilmore, 2006). Although highly conserved, and despite structural data for human Ost4 (Gayen and Kang, 2011), not much was known about the function of mammalian OST4 until a recent study (Dumax-Vorzet et al., 2013). It was shown that mammalian OST4 was important for stability of the OST complex and knockdown resulted in a mild defect in the N-glycosylation of an endogenous HeLa glycoprotein. Also, in yeast, point mutations in the transmembrane part of OST4p disrupt the OST4, OST3, and STT3 interaction and cause growth defects (Kim et al., 2000, 2003).

STT3 is the largest protein subunit of the OST complex (Spirig et al., 1997). About 10 years ago, data providing evidence about the catalytic role of STT3 in OST complexes was presented (Kelleher et al., 2003; Nilsson et al., 2003). The N-terminal regions of STT3 proteins contain a hydrophobic membrane-anchor domain of 11 or 13 predicted transmembrane helices followed by a soluble, catalytic C-terminal domain which extends into the ER lumen(Kelleher and Gilmore, 2006). In mammals, two isoforms of STT3 (STT3A and STT3B) are encoded in the genome.However, the two proteinsare not incorporated into the same complex, insteadtwo different OSTs, with either STT3A or STT3B, are expressed (Kelleher et al., 2003). The two OST complexes are co-expressed in several tissues, but they vary inexpression levels. Human STT3A and STT3B proteins share 59% amino acid sequence identity and STT3B is slightly larger than STT3A(Kelleher et al., 2003). Furthermore, OST complexes with STT3B as the catalytic subunit have been reported to be more active than the STT3A complex form (Kelleher et al., 2003; Ruiz-Canada et al., 2009). Additionally, STT3A and STT3B have been shown to have specific acceptor substrate selectivity. In 2007, Wilson and High analysed the glycosylation state of specific membrane proteins subsequent to siRNA-mediated STT3A/STT3B knockdown in HeLa cells. When depleting STT3A, loss of both OST isoforms was seen, accompanied by the abolition of glycosylation. In contrast, depletion of STT3B did not abolish glycosylation completely, instead hypoglycosylation was seen on certain membrane proteins(Wilson and High, 2007). Although a more recent study presented contradicting data, suggesting that depletion of STT3A or STT3B increases the expression level of the non-affected protein, the same study provided insight to the order of glycosylation of STT3A and STT3B(Ruiz-Canada et al., 2009). OST complexes containing STT3A were shown to catalyse co-translational transfer of the precursor oligosaccharide, whereas STT3B complexes catalysed post-translational glycosylation of remaining sequons (Ruiz-Canada et al.,

2009). This study also suggested distinct enzymatic properties in that STT3B complexes (unlike the STT3A equivalent) could mediate glycosylation of asequon found adjacent to the C-terminus of a secreted protein(Ruiz-Canada et al., 2009). Lastly, a recent study demonstrated that N-X-S sequons withpositively charged, small hydrophobic andpolar side chains in the X position are efficiently co-translationally glycosylated by STT3A OST. However bulky, hydrophobic or negatively charged amino acids in the X position were frequently skipped during co-translational glycosylation by *Tb*STT3A OST (Malaby and Kobertz, 2014).

The two mammalian proteins ,MagT1/IAP and TUSC3/N33, were initially suggested to be homologous to the yeast Ost3 and Ost6 proteins(Kelleher et al., 2003; MacGrogan et al., 1996). Despite low sequence identity (20%) between TUSC3 or MagT1 and yeast Ost3/Ost6, the four proteins share a lumenal thioredoxin domain (Fetrow et al., 2001; Kelleher and Gilmore, 2006; Mohorko et al., 2014). Furthermore, mutation of the CXXC motifs in Ost3 or Ost6 decreases glycan occupancy on specific sequons in yeast glycoproteins (Schulz et al., 2009). Recently, it has been proposed that MagT1/IAP exclusively assemble with the STT3B OST complex and that it can enable glycosylation of sequons that were not glycosylated by the STT3A complex and that are close to cysteine amino acids. This is proposed to be achieved by establishing a transient disulfide bridge between MagT1 and the glycoprotein substrate and this mechanism helps reduce the competition between disulfide bond formation and N-glycosylation (Cherepanova et al., 2014; Mohorko et al., 2014).

The most recent discovered subunits that associate with the OST complex are KCP2 and DC2. DC2 shows weak homology to the C-terminal half of Ost3 and Ost6 in yeast. However, no domain homology was found for KCP2. A recent study revealed that KCP2 is important for stability of the STT3A-containing OST complex and that it plays

a part in substrate selectivity by this complex(Roboti and High, 2012a; Shibatani et al., 2005).

As previously mentioned, N-linked glycosylation is also found in bacteria and archaea, however, OSTs in these organisms are not protein heterologous complexes. Instead, they are composed by single proteins that are homologues to STT3, called PglB in bacteria and AglB in archaea (Chaban et al., 2006; Dell et al., 2010; Guan et al., 2010; Jervis et al., 2010; Kelleher et al., 2007; Kelly et al., 2009; Schwarz et al., 2011). Archaea displays a great diversity in glycosylation pathways, by using many different sugar building blocks and requiring a wider variety of donor substrates than found in eukaryotes. Bacteria exhibit a more limited N-glycan range than archaea, and N-linked glycosylation is restricted to a relatively small number of species (Guan et al., 2010; Kelleher et al., 2007). The first complete OST crystal structure that was solved was from the bacteria *Campylobacter lari*, giving many insights about the architecture of this class of enzymes together with possible modes of mechanisms (Lizak et al., 2011). One long loop (EL5), which connects two transmembrane helices, was reported to be only partially ordered, with 25 residues being disordered in the electron density map. Based on the proposed mobility of the EL5 loop, a three state mechanism was proposed for PglB catalysed glycosylation (Figure 1.6). Upon binding the acceptor peptide, the C-terminal end of EL5 is ordered and pins the sequon-containing peptide to the periplasmic domain, thus restricting the mobility of the peptide. Once the LLO is bound by PglB, a nucleophilic attack of the activated amide nitrogen occurs and the nascent protein becomes glycosylated. Subsequently, the attached oligosaccharideis closely pressed against PglB, causing steric tension which will initiate disengagement of EL5 and allow release of the glycosylated protein. However, there are no data as to which event occurs first, the binding of the acceptor peptide or the binding of the LLO (Lizak et al., 2011).

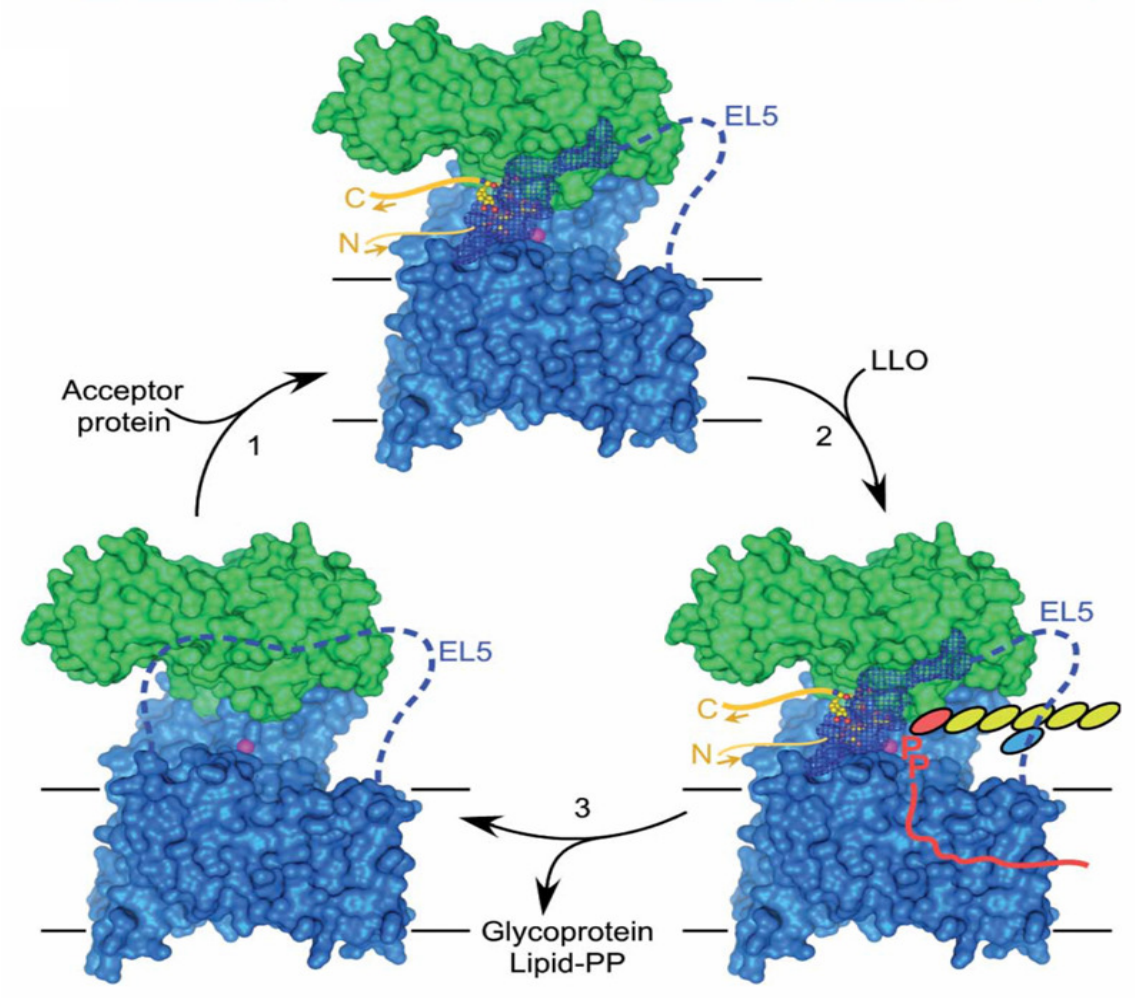


Figure 1.6. Proposed mechanism for N-glycosylation in *C. lari*. 1) In the first stage, the EL5 loop is going from being disordered, generating a possibility to bind the acceptor protein, to having the C-terminal of the EL5 loop ordered and bound to the acceptor peptide. Thus, the C-terminal part of the EL5 loop pins the nascent protein against the periplasmic domain, hence limiting the mobility of the peptide. 2) In the second stage, the LLO binds the other cleft of the *C. lari* OST, accommodating proximity of the two substrates to generate glycosylation. 3) In the third and last step, the loop is disordered again, enabling the release of glycosylated sequon and lipid-pyrophosphate (PP)(Lizak et al., 2011)

A more recent study supported the view of the EL5 loop being part of the binding process of the acceptor peptide, but the study also elaborated the role of the N-terminus of the loop (Lizak et al., 2014). The first 25 residues of the EL5 loop were not visible in the electron density map, however, a conserved tyr/phe motif was shown to be essential for catalysing the glycan transfer. The motif is believed to interact with the LLO, possibly by forcing the reducing end of the sugar into a preferred, reactive conformation (Lizak et al., 2014).

It is important to note that alternative N-glycosylation pathways have been discovered in a few bacteria (Choi et al., 2010; Grass et al., 2003, 2010; Gross et al., 2008). These reactions take place in the cytoplasm, where monosaccharides from activated sugar donors are transferred to an asparagine in the same sequon used by OST, N-X-S/T. The enzyme catalysing this transfer, N-glycosyltransferase (NGT), does not belong to the same structural class of enzyme as OST, implying that the same sequon specificity has emerged twice in evolution and has resulted in parallel enzymatic activity (Choi et al., 2010; Grass et al., 2003, 2010; Gross et al., 2008).

1.4 N-linked glycosylation and oligosaccharyltransferase activities in *T. brucei*

1.4.1 Glycoproteins in *T. brucei*

T. brucei has several known glycoproteins (VSG, ISG65 and ISG75, Fla1, the transferrin receptor, p67, tGLP1, GPIdeAc and *Tb*MBAP1)and many of them are essential for cell survival (Cross, 1975a; Engstler et al., 2005; Güther et al., 2003; Kelley et al., 1999; Lingnau et al., 1999; Mehlert, 1998; Mehlert et al., 2012; Nozaki, 1996; Steverding et al., 1994; Wooten et al., 1990; Ziegelbauer and Overath, 1992). The

different types of oligosaccharides that can be found on *T. brucei* proteins are summarised in Figure 1.7.

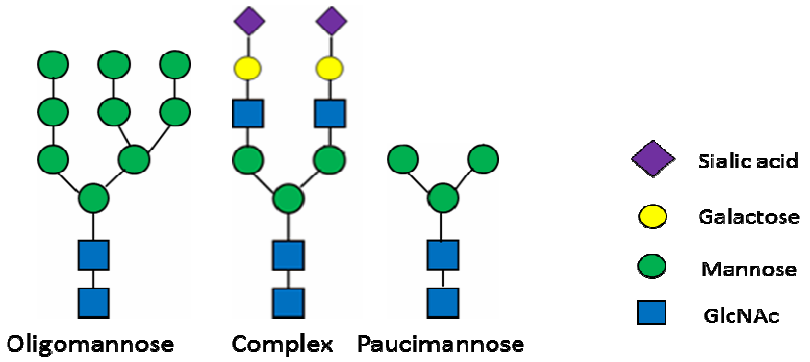


Figure 1.7. Typical samples of the different types of oligosaccharides found bound to the asparagine in the N-X-S/T sequon of *T.brucei*glycoproteins.

The VSGs form a densely packed layer of dimers on the cell surface which are anchored in the membrane byglycosylphosphatidylinositol (GPI)(Ferguson et al., 1988). The N-linked glycan structure of a fewvariants of *T.brucei* VSGs have been solved,showing that each VSG variant contains one, two, or three N-linked oligosaccharides per monomer(Mehlert et al., 2002, 2010; Strang et al., 1993; Wooten et al., 1990; Zamze et al., 1991). VSG221 is found on the Lister 427 *T. brucei* strain used for the work carried out in this thesis and this VSG variant contains two N-glycosylation sites, one at Asn263 and one at Asn428, occupied by different types of oligosaccharides. The oligosaccharides covalently attached to the asparagine in position 263, consist of small biantennary complex and paucimannose oligosaccharides.The asparagine at position 428 is occupied mostly by oligomannose glycans.(Zamze et al., 1991). The transferrin receptor is another glycoprotein in *T. brucei* and it is localised in the flagellar pocket where it is anchored to the membrane via a GPI(Steverding et al., 1994). The receptor is a heterodimer consisting of the expression site associated gene 6 and 7 (ESAG6 and

ESAG7) and has recently been reported to have eight N-glycosylation sequons occupied with either oligomannose or paucimannose glycans (Mehlert et al., 2012).

1.4.2 Oligosaccharyltransferase in *T. brucei*

Just like in bacteria and archaea, OST in *T. brucei* is suggested to comprise of a single protein, the catalytic subunit STT3. No other OST subunit homologues were found in in the *T. brucei*genome, however, three STT3 genes (*Tb*STT3A, *Tb*STT3B and *Tb*STT3C) were found in a tandem array on the fifth chromosome (Figure 1.8) (Berriman, 2005; Kelleher and Gilmore, 2006). It is possible that *T. brucei* diversified the single subunit OST by gene duplication, in a similar manner to the related kinetoplastid parasite *Leishmania major*(Hese et al., 2009; Nasab et al., 2008).

Figure 1.8.STT3 locus in *T. brucei.Tb*STT3A shares 79 % sequence identity with *Tb*STT3B and *Tb*STT3C, whereas *Tb*STT3B and *Tb*STT3C are more closely related with 95 % sequence identity. There are identical UTRs surrounding the three genes, highlighted in pink and red.

The encoded protein sequences of the three genes are highly conserved, with *Tb*STT3B and *Tb*STT3C showing 95 % identity, whereas *Tb*STT3A share 79 % identity with them both. The three genes are found in a highly repetitive segment of the genome and are surrounded by identical untranslated regions (UTRs), highlighted in red and pink in Figure 1.8. Both *Tb*STT3A and *Tb*STT3B were shown to be expressed at the mRNA

level in the life cycle stages we culture in the laboratory (bloodstream form and procyclic form) and a recent quantitative proteomics study showed that both are more highly expressed at the protein level in bloodstream form parasites, 8-fold for *Tb*STT3A and 1.5-fold for *Tb*STT3B. *Tb*STT3C, on the other hand, were not detected at the mRNA or protein level and has been suggested to possibly play a role in different life stages (Izquierdo et al., 2009a; Urbaniak et al., 2013).

Already in 1988, Bangs and colleagues suggested selective mechanisms for glycosylation in *T. brucei* based on the two specific types of glycan found bound to certain sequons on a VSG molecule (Bangs et al., 1988). This assumption has later been supported by different studies and Ferguson and colleagues have provided data about LLO specificity for *Tb*STT3A and *Tb*STT3B (Izquierdo et al., 2009a, 2012; Jones et al., 2005; Manthri et al., 2008). The biantennary $Man_5GlcNAc_2$ oligosaccharide is recognised and transferred by *Tb*STT3A whereas the triantennary $Man_9GlcNAc_2$ oligosaccharide is the preferred oligosaccharide choice of *Tb*STT3B (Figure 1.9 and 1.10). Having two distinct precursor oligosaccharides, transferred by two different OSTs, is unique for *T. brucei*. Mammalian and yeast OST complexes are selective for the oligosaccharide end product $Glc_3Man_9GlcNAc_2$ whilst protists and fungi exhibit either little specificity for the donor LLO structure or are selective to the end LLO product (Kelleher et al., 2007; Samuelson et al., 2005). In *T. brucei*, the $Man_5GlcNAc_2$-PP-Dol molecule is the most abundant lipid-linked oligosaccharide in both bloodstream and procyclic life stages. Studies have shown that if *Tb*STT3A is depleted, *Tb*STT3B is able to transfer trianntennary $Man_9GlcNAc_2$ oligosaccharides to some of the sequons normally acted upon by *Tb*STT3A. However, with *Tb*STT3A absent, significant hypoglycosylation occurs. When *Tb*STT3B is depleted, *Tb*STT3A is barely able to glycosylate sequons normally acted upon by *Tb*STT3B and if it does, it exclusively transfers $Man_5GlcNAc_2$. Therefore, similar to *Tb*STT3B depletion, loss of *Tb*STT3A

also leads to hypoglycosylation in *T. brucei*(Izquierdo et al., 2009a). Additionally, a more recent study showed that *Tb*STT3A could transfer $Man_7GlcNAc_2$with approximately the same efficiency as $Man_5GlcNAc_2$ whereas, although *Tb*STT3B could use $Man_7GlcNAc_2$with approximately the same efficiency as $Man_5GlcNAc_2$, both donors were much less efficient than the natural oligosaccharide $Man_9GlcNAc_2$. Hence, it was suggested that *Tb*STT3A transfer is reliant on the LLO not having a C-branch, whereas *Tb*STT3B seems to require the C-branch for efficient LLO transfer (Izquierdo et al., 2012) (Figure 1.9).

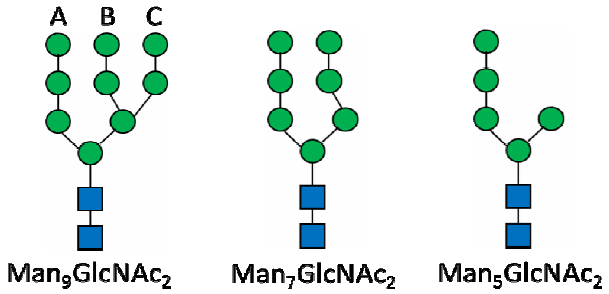


Figure 1.9. Structures of the three oligosaccharides $Man_9GlcNAc_2$, $Man_7GlcNAc_2$ and $Man_5GlcNAc_2$. The A/B/C-branches are highlighted in $Man_9GlcNAc_2$. *Tb*STT3A recognises $Man_7GlcNAc_2$ equally well as the native donor $Man_5GlcNAc_2$, but cannot bind and/or transfer the triantennary $Man_9GlcNAc_2$. *Tb*STT3B can recognise and transfer both $Man_7GlcNAc_2$ and $Man_5GlcNAc_2$, though significantly less efficient than $Man_9GlcNAc_2$. Therefore *Tb*STT3B prefers oligosaccharides with the C-branch, while *Tb*STT3A glycan transfer is inhibited by the same (Izquierdo et al., 2012).

The related kinetoplastid parasites *Leishmania major*and *Trypanosoma cruzi* have also been reported to only encode the catalytic STT3 subunit, but similar to *T.brucei*, *Leishmania major* have several copies of this gene and probably express STT3 enzymes

with distinct but overlapping peptide substrate specificities. Some of the single subunit OSTs from these parasites have been shown to complement STT3 deletion in yeast (Castro et al., 2006; Nasab et al., 2008). Similarly, *Tb*STT3B and *Tb*STT3C were able to support cell viability in yeast following depletion of the endogenous STT3. Interestingly, *Tb*STT3A could not (Izquierdo et al., 2012).

When analysing the acceptor peptide specificity of *Tb*STT3A and *Tb*STT3B in bloodstream form parasites, overlapping but distinct selectivity was observed (Izquierdo et al., 2009a). By examining the sequons and surrounding peptide sequences (XXXXX**NXS/T**XXXXX) where the attached glycan type (biantennary transferred by *Tb*STT3A or triantennary transferred by *Tb*STT3B) were known, a pattern for *Tb*STT3A catalysed transfer was seen. These sequons were often surrounded by acidic amino acids and thus the sequon peptide had a relatively low pI value. However, *Tb*STT3B, which catalyses post-translational oligosaccharide transfer, was shown to have no specific preference concerning the pI value of the sequon-flanking residues. Instead, *Tb*STT3B, having a broad sequon peptide specificity, was proposed to glycosylate sequons that had been overlooked by *Tb*STT3A (Figure 1.10). This distinct but overlapping dual glycosylation has been suggested to produce a high efficiency of oligosaccharide transfer in *T. brucei*(Izquierdo et al., 2009a). Furthermore, the acceptor peptide specificity of *Tb*STT3C was tested in a yeast model, highlighting a tendency to glycosylate sequons in an acidic environment, similar to *Tb*STT3A (Izquierdo et al., 2009a).

1.4.3 N-glycan processing in *T. brucei*

Oligosaccharide processing in *T. brucei* occurs both in the ER and in the Golgi apparatus. On the lumenal side of the ER, proteins like UDP-Glc:glycoprotein glucosyltransferase (UGGT), α-glucosidase II (GII) and calreticulin act upon the glycan in order to aid folding of nascent proteins (Izquierdo et al., 2009b). Furthermore, many mature glycoproteins in the bloodstream form are equipped with complex/paucimannoseand/or oligomannose oligosaccharides, for example (as aforementioned), the complex and paucimannose glycans found attached to the asparagine at position 263 of VSG221, the paucimannose glycans attached to the asparagine at position 94 on the ESAG6 monomer of the transferrin receptor and the oligomannose glycans found at position 10 and 344 on the same monomer of the receptor and on position 428 of VSG221 (Jones et al., 2005; Mehlert et al., 2012; Zamze et al., 1991).

The two precursor oligosaccharides used in *T. brucei*, $Man_5GlcNAc_2$ and $Man_9GlcNAc_2$, can only be processed into specific types of glycans. The biantennary $Man_5GlcNAc_2$ is the only route to complex or paucimannose oligosaccharides, whereas $Man_9GlcNAc_2$ will only be processed into oligomannose glycans. Presumably, these separate routes have evolved from *T. brucei* lacking a Golgi α-mannosidase II gene, such that the oligomannose type of glycan can only be processed as far as triantennary $Man_5GlcNAc_2$. Thus, the massive poly-*N*-acetyllactosamine structures that have been reported in the bloodstream form of the parasite, originates from *Tb*STT3A transfer of the biantennary $Man_5GlcNAc_2$(Atrih et al., 2005; Manthri et al., 2008).

The steps to produce these complex glycans occur in the Golgi and are catalysed by mannosidase and glycosyltransferases, and recently two glycosyltransferases, *Tb*GT8 and TbGT11, were found to be involved in the construction of complex glycans in both

bloodstream (*Tb*GT8 and *Tb*GT11) and procyclic form (*Tb*GT8) (Damerow et al., 2014; Nakanishi et al., 2014).

Following RNAi induced knockdown of *Tb*STT3A and *Tb*STT3B, knowledge about the two single-subunit OST essentiality in bloodstream form parasites was gathered (Izquierdo et al., 2009a). Both proteins were found to be non-essential *in vitro* when knocked down individually, however, simultaneous knockdown caused a cessation in cell growth and proved that protein N-glycosylation is essential in *T.brucei*. Interestingly, individual knockdown of either *Tb*STT3A or *Tb*STT3B was found to abolish the ability of the parasite to infect mice, highlighting that both *Tb*STT3A and *Tb*STT3B glycosylation mechanisms are independently essential *in vivo*(Izquierdo et al., 2009a).

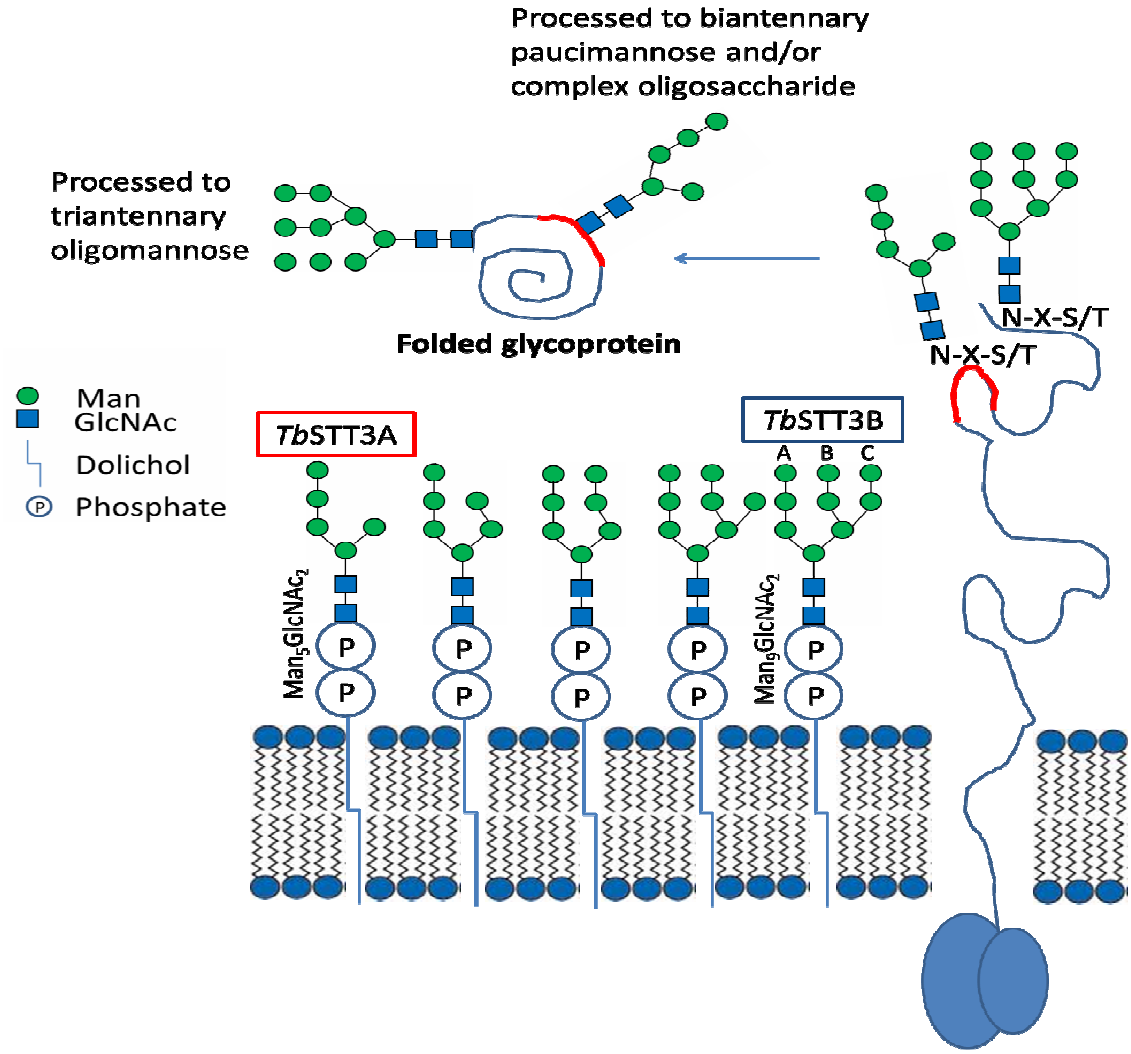


Figure 1.10. Schematic view of lipid-linked oligosaccharide and sequon specificity in *T. brucei*.$Man_5GlcNAc_2$ is recognised and co-translationally transferred to selected sequons surrounded by acidic amino acids (highlighted in red in nascent protein)by *Tb*STT3A. $Man_9GlcNAc_2$ is recognised and post-translationally transferred to any remaining, sterically accessible sequons (independent of surrounding pH) by *Tb*STT3B. *Tb*STT3A transfer is dependent on the LLO not having a C-branch, whereas *Tb*STT3B require the C-branch for efficient LLO transfer (Izquierdo et al., 2009a, 2012).

Figure 1.10.(continued)The biantennary $Man_5GlcNAc_2$ is the only route to paucimannose and complex oligosaccharides, while the triantennary $Man_9GlcNAc_2$ can only be processed to oligomannose glycans because of the lack of a Golgi α-mannosidase II gene (Manthri et al., 2008).

1.5 Aim of the thesis

The project wasdivided in three component parts:

1. Generation of *Tb*STT3A gene replacement mutants and analysis of their biochemical and biological phenotype.
2. *Tb*OST localisation studies and search for novel *Tb*STT3A binding partners.
3. Testing the sequon specificity of *Tb*STT3A and *Tb*STT3B *in vivo.*

1. The aim of this project is to investigate the essentiality and function of *Tb*STT3A and the families of N-glycans they are ultimately responsible for in mature *T. brucei* glycoproteins. To address these goals, gene replacement experiments in order to generate null and/or conditional null mutants of *Tb*STT3A, will be performed.
2. For the localisation and putative *Tb*OST subunits studies, endogenous *Tb*STT3A will be *in situ* tagged. This tag enables both immunolocalisation and immunoprecipitation experiments which will help elucidate where in the cell *Tb*STT3A is localised and if it has any binding partners.
3. Using a glycosylation reporter protein, N-linked sequon specificity of *Tb*STT3A and *Tb*STT3B will be investigated in order to investigate the role of flanking residues involvement in N-glycosylation.

2 – Materials and Methods

2.1 Materials

All reagents, enzymes and chemicals used in this study were of the highest grade and purity available from commercial companies.

2.2 Parasite culture

2.2.1 *T. brucei*cell lines and media

Strain 427 of *T. brucei* bloodstream form cells were grown in HMI-9t medium supplemented with glutamax at 37°C with 5% CO_2. HMI-9t is similar to the HMI-9 medium previously described earlier (Hirumi and Hirumi, 1989), but thioglycerol has replaced β-mercaptoethanol. Parasites were grown in tissue culture flasks with filter lids to a maximum density of $5x10^6$ cells/ml. The cell line which was used in the experiments was the single marker (sm) cell line, which expresses a tetracycline repressor protein (TetR) and a T7 polymerase (Wirtz *et al*., 1999) under neomycin (G418) selection. Table 2.1 gives the antibiotic concentration used to maintain selection.

Table 2.1. Antibiotics used in *T. brucei* cell culture and genetic manipulation

Antibiotic	**Resistance gene**	**Concentration µg/ml**
Neomycin (G418)	Neomycin phosphotransferase	2.5
Hygromycin	Hygromycin phosphotransferase	4
Phleomycin	Phleomycin resistance protein	0.1
Puromycin	Puromycin acetyltransferase	2.5

Blasticiden	Blasticidin deaminase	2.5
Tetracycline		0.5

2.2.2 General *T. brucei* cell culture protocols

2.2.2.1 Cell density

Cells were counted with the CASY® Cell Counter + Analyser system, which measures the cell density in the culture. Alternatively, cells were counted in a haemocytometer, diluted 1:10 in 1x trypanosome dilution buffer (TDB)(Table 2.2).

Table 2.2. 5x Trypanosome Dilution Buffer (TDB)

Chemical substances	**Concentration (M)**
KCl	0.025
NaCl	0.4
$MgSO_4 \bullet H_2O$	0.005
Na_2HPO_4	0.1
$NaH_2PO_4 \bullet H_2O$	0.01
Glucose	0.1

2.2.2.2 Generation of stabilates

Cells were grown to mid log phase (1-3x10^6 cells/ml), centrifuged (800 g, 10 min, 4°C on 4K15 Sigma centrifuge) and resuspended in half of the starting volume HMI-9t medium, containing 10% glycerol and no antibiotics. Aliquots of 0.5 ml were frozen in 2 ml Cryovials (Nalgene).

2.2.2.3 T. brucei crude cell lysate

Cells were grown to mid log phase density and 3x10^7 cells were harvested by 10 min centrifugation (4K15 Sigma Centrifuge) at 4°C. Cells were washed twice in TDB buffer at 4°C. After the last wash, cells were transferred to a 1.5 ml tube and pelleted for 30s at 11000 g (Centrifuge 5417R, Eppendorf). The cells were then resuspended in 1xTDB

buffer and lysed by addition of an equal volume of 2% SDS sample buffer (Invitrogen).The cell lysate was heated to 95°C for 10 min before loading on a gradient SDS-PAGE gel (Invitrogen).

2.2.2.4 T. brucei bloodstream form electroporation

Cells were grown to mid log phase density (1-3x10^6 cells/ml),centrifuged (800 g, 10 min, 20°C) and resuspended in Amaxa® Human T-Cell Nucleofector solution containing supplement (Lonza) so that the final concentration was 1x10^8 cells/ml. 100 µl was transferred to cuvettes (Amaxa certified, included in the Nucleofector kit) containing 1 µg of plasmid DNA each.One cuvette contained no DNA and was used as a negative control. The cells were electroporated using an Amaxa Nucleofector® II, program X-001 (Burkard *et al*., 2007). Subsequent to electroporation, the cells from each cuvette were transferred in a flask with 50 ml of HMI-9t containing all necessary antibiotics except the selection antibiotic for the electroporated construct. The cells were left to recover for 6-8 h at 37°C before the selection antibiotic was added. When the control cells, electroporated without any DNA construct and hence no antibiotic resistance gene, were dead, resistant clones were established in culture, stabilated and their genotype analysed by Southern blot.

2.3 General molecular biology

2.3.1 Primers

Primers were designed to contain ~20 bp of complementary sequence, preferably with a guanine (G) or cytosine (C) at their 3’ end and an approximate melting temperature (Tm) of 68°C. The melting temperaturewas calculated by the following formula: Tm=

(G+C)x 4°C + (A+T)x 2°C.Only the nucleotides in the complementary sequence of the primer were considered in the calculations. All primers were synthesized by the Oligonucleotide Synthesis Service, University of Dundee. Primers used are listed in table 2.3.

Table 2.3. List of primers

Function	**Primers**	**Sequence**
Knock out	5' UTR Fw	ATAAGTATGCGGCCGCAAATGTCCATCGACTCTCATTCC
Knock out	5' UTR Rev	GTTTAAACTTACGGACCGTCAAGCTTTAAGCAGAAAGAAAC GCCCAAAG
Knock out	3' UTR Fw	GACGGTCCGTAAGTTTAAACGGATCCGTCGAAAACCTGTATT TTCAGAGC
Knock out	3' UTR Rev	ATAAGTAAGCGGCCGCCTATTCCTTTGCCCTCGGACG
Ectopic copy	STT3A Fw	ATAAGTATAAGCTTATGACGAAAGGTGGGAAAGTAGC
Ectopic copy	STT3A Rev	ATAAGTAAGGATCCTAGCGAAGCGGTTCCATTACGAA
In situ tagging	STT3A 4YH 5' Fw	AGTATAAAGGTACCGTTACCCTCTTTGGGGGTGCG
In situ tagging	STT3A 4YH 5' Rev	TTACTTATGGGCCCTTCGTAATGGAACCGCTTCGCT
In situ tagging	STT3A 4YH 3' Fw	ATAATGAAGCGGCCGCTGACCACATCGTTTCAATCGCC
In situ tagging	STT3A 4YH 3' Rev	TTACTTATCCACCGCCGTGGGCGACGAACGTAGTGGATGAC
In situ tagging	STT3A 4H 5' Fw	ATAAGTATCTCGAGCAAGTTTGCTTGCCCCGTTCG
In situ tagging	STT3A 4H	ATAAGTAACTCGAGCTCGCTCTGAAAATACAGGTTTTCGACT

	5’ Rev	TCGTAATG GAACCGCTTCGCT
In situ tagging	STT3A 4H 3’ Fw	ATAAGTATGGATCCCCACATCGTTTCAATCGCCGC
In situ tagging	STT3A 4H 3’ Rev	ATAAGTAAGGATCCACTCACAATCGTGCTTACAGCC
Southern blot probe	STT3A probe Fw	CCCTCTTTGGGGGTGCGGTG
Southern blot probe	STT3A probe Rev	ATTGGGTAGATCAGTCACGGCA
Southern blot probe	STT3A/B/C probe Fw	TTGGCCGACCGGTGGGCAC
Southern blot probe	STT3A/B/C probe Rev	GCGTCCACTCCACTGGCGGT
PCR knock out control	Upstream 5’ UTR Fw	GCAGAAGCGACAACTAGACCTC
PCR knock out control	3xHA rev	CTATGCGTAATCGGGCACATCGT
PCR knock out control	BSD Rev	TTGTAGCCGTTGCTCTTTCAATG
BipN reporter protein	BipN Fw	ATAAGTATAAGCTTATGTCGAGGATGTGGC
BipN reporter protein	BipN Rev	ATAAGTAAAATTAATTCCTAGGCCCGCCAACCTCGCTTTCAC
Glycosequon assay	STT3B 1 Fw	CTAGGGCGGCGGCGGATTTCAATAGGCGGAATCGCTCGGAT GCTTACTATAGAC
Glycosequon	STT3B 1	AATTGTCTATAGTAAGCATCCGAGCGATTCCGCCTATTGAAA

assay	Rev	TCCGCCGCCGCC
Glycosequon assay	STT3A/B 1 Fw	CTAGGGCGGCGGCGTATTTGTGGCTAAGAAACTCGACACCG GAAGATGCCAGGC
Glycosequon assay	STT3A/B 1 Rev	AATTGCCTGGCATCTTCCGGTGTCGAGTTTCTTAGCCACAAA TACGCCGCCGCC
Glycosequon assay	STT3A/B 2 Fw	CTAGGGCGGCGGCGATTACAGGGATCGGTAACAGGACGACC CTGGCGGACGGTC
Glycosequon assay	STT3A/B 2 Rev	AATTGACCGTCCGCCAGGGTCGTCCTGTTACCGATCCCTGTA ATCGCCGCCGCC
Glycosequon assay	STT3A/B 3 Fw	CTAGGGCGGCGGCGATATATAAGGTGATGAATGTTTCGGAG GAATCGAAAGCAC
Glycosequon assay	STT3A/B 3 Rev	AATTGTGCTTTCGATTCCTCCGAAACATTCATCACCTTATATA TCGCCGCCGCC
Glycosequon assay	Transferrin 1 Fw	CTAGGGCGGCGGCGGAAAGAAATGCACTGAACGCTACTGCT GCGAATAAAGTGC
Glycosequon assay	Transferrin 1 Rev	AATTGCACTTTATTCGCAGCAGTAGCGTTCAGTGCATTTCTTT CCGCCGCCGCC
Glycosequon assay	Transferrin 2 Fw	CTAGGGCGGCGGCGCTTGAAGAGATGAGAAACGCCTCAGCG CTCGCAGCCGCGGCGC
Glycosequon assay	Transferrin 2 Rev	AATTGCGCCGCGGCTGCGAGCGCTGAGGCGTTTCTCATCTCT TCAAGCGCCGCCGCC
Glycosequon assay	Transferrin 3 Fw	CTAGGGCGGCGGCGACGATACTTAAGTCGAATTATACTGCA GAACCGGTGAGGC
Glycosequon assay	Transferrin 3 Rev	AATTGCCTCACCGGTTCTGCAGTATAATTCGACTTAAGTATC GTCGCCGCCGCC
Glycosequon assay	Transferrin 4 Fw	CTAGGGCGGCGGCGGAACGAAACGCTCTGAACGCTACCGCC GCTAATAAAGTAC
Glycosequon assay	Transferrin 4 Rev	AATTGTACTTTATTAGCGGCGGTAGCGTTCAGAGCGTTTCGT TCCGCCGCCGCC
Glycosequon assay	Transferrin 5 Fw	CTAGGGCGGCGGCGTCTCCCACCAGGCACAACCTAACCTGG GGAGGTGGTGTTC

Glycosequon assay	Transferrin 5 Rev	AATTGAACACCACCTCCCCAGGTTAGGTTGTGCCTGGTGGGAGACGCCGCCGCC
Glycosequon assay	Transferrin 6 Fw	CTAGGGCGGCGGCGTTTGGGTCATATCAGAACGGTTCAATGTATGTAGAGGGCC
Glycosequon assay	Transferrin 6 Rev	AATTGGCCCTCTACATACATTGAACCGTTCTGATATGACCCAAACGCCGCCGCC
Glycosequon assay	Transferrin 7 Fw	CTAGGGCGGCGGCGGTGTGCCTCAATCGCAACTTCACTTGGGGCGGGGGTGTCC
Glycosequon assay	Transferrin 7 Rev	AATTGGACACCCCCGCCCCAAGTGAAGTTGCGATTGAGGCACACCGCCGCCGCC
Glycosequon assay	Alanine control Fw	CTAGGGCGGCGGCGGCGGCGGCGGCGGCGAACGCGACGGCGGCGGCGGCGGCGC
Glycosequon assay	Alanine control Rev	AATTGCGCCGCCGCCGCCGCCGTCGCGTTCGCCGCCGCCGCCGCCGCCGCCGCC
Glycosequon assay	D1 Fw	CTAGGGCGGCGGCGGATGCTGCGGCCGCAAATGCAACAGCCGCAGCAGCAGCGC
Glycosequon assay	D1 Rev	AATTGCGCTGCTGCTGCGGCTGTTGCATTTGCGGCCGCAGCATCCGCCGCCGCC
Glycosequon assay	D2 Fw	CTAGGGCGGCGGCGGCCGATGCAGCCGCGAATGCGACGGCCGCCGCTGCAGCCC
Glycosequon assay	D2 Rev	AATTGGGCTGCAGCGGCGGCCGTCGCATTCGCGGCTGCATCGGCCGCCGCCGCC
Glycosequon assay	D3 Fw	CTAGGGCGGCGGCGGCAGCCGATGCGGCCAATGCTACGGCCGCTGCAGCCGCAC
Glycosequon assay	D3 Rev	AATTGTGCGGCTGCAGCGGCCGTAGCATTGGCCGCATCGGCTGCCGCCGCCGCC
Glycosequon assay	D4 Fw	CTAGGGCGGCGGCGGCAGCTGCAGATGCGAACGCCACGGCCGCGGCTGCTGCCC
Glycosequon assay	D4 Rev	AATTGGGCAGCAGCCGCGGCCGTGGCGTTCGCATCTGCAGCTGCCGCCGCCGCC
Glycosequon	D5 Fw	CTAGGGCGGCGGCGGCTGCCGCCGCTGACAATGCAACAGCC

assay		GCGGCCGCGGCAC
Glycosequon assay	D5 Rev	AATTGTGCCGCGGCCGCGGCTGTTGCATTGTCAGCGGCGGCAGCCGCCGCCGCC
Glycosequon assay	D6 Fw	CTAGGGCGGCGGCGGCTGCAGCTGCTGCGAATGACACCGCAGCTGCGGCAGCCC
Glycosequon assay	D6 Rev	AATTGGGCTGCCGCAGCTGCGGTGTCATTCGCAGCAGCTGCAGCCGCCGCCGCC
Glycosequon assay	D7 Fw	CTAGGGCGGCGGCGGCGGCCGCGGCCGCAAATGCTACAGATGCAGCGGCCGCAC
Glycosequon assay	D7 Rev	AATTGTGCGGCCGCTGCATCTGTAGCATTTGCGGCCGCGGCCGCCGCCGCCGCC
Glycosequon assay	D8 Fw	CTAGGGCGGCGGCGGCGGCCGCCGCAGCAAACGCGACCGCGGATGCCGCCGCAC
Glycosequon assay	D8 Rev	AATTGTGCGGCGGCATCCGCGGTCGCGTTTGCTGCGGCGGCCGCCGCCGCCGCC
Glycosequon assay	D9 Fw	CTAGGGCGGCGGCGGCTGCGGCAGCAGCGAACGCGACAGCGGCAGATGCAGCGC
Glycosequon assay	D9 Rev	AATTGCGCTGCATCTGCCGCTGTCGCGTTCGCTGCTGCCGCAGCCGCCGCCGCC
Glycosequon assay	D10 Fw	CTAGGGCGGCGGCGGCAGCGGCAGCGGCGAACGCGACGGCTGCTGCCGATGCAC
Glycosequon assay	D10 Rev	AATTGTGCATCGGCAGCAGCCGTCGCGTTCGCCGCTGCCGCTGCCGCCGCCGCC
Glycosequon assay	D11 Fw	CTAGGGCGGCGGCGGCCGCCGCGGCTGCTAACGCGACCGCGGCGGCGGCAGACC
Glycosequon assay	D11 Rev	AATTGGTCTGCCGCCGCCGCGGTCGCGTTAGCAGCCGCGGCGGCCGCCGCCGCC
Point mutation	STT3B H397R Fw	GTGGATTCTGTGGCTGAGCACCGTCCGGCGTCGAATGATGATTTC
Point	STT3B	GAAATCATCATTCGACGCCGGACGGTGCTCAGCCACAGAAT

mutation	H397R Rev	CCAC
Point mutation	STT3C R397H Fw	GTGGATTCTGTGGCTGAGCATCACCCGACGACTGCCGGGGC CTTC
Point mutation	STT3C R397H Rev	GAAGGCCCCGGCAGTCGTCGGGTGATGCTCAGCCACAGAAT CCAC
Point mutation	STT3C R397D Fw	GTGGATTCTGTGGCTGAGCATGACCCGACGACTGCCGGGGC CTTC
Point mutation	STT3C R397D Rev	GAAGGCCCCGGCAGTCGTCGGGTCATGCTCAGCCACAGAAT CCAC
Point mutation	STT3C R397Q Fw	GTGGATTCTGTGGCTGAGCATCAGCCGACGACTGCCGGGGC CTTC
Point mutation	STT3C R397Q Rev	GAAGGCCCCGGCAGTCGTCGGCTGATGCTCAGCCACAGAAT CCAC

2.3.2 List of plasmid vectors

Table 2.4. List of plasmids

Plasmid	**Use**	**Source**
pGEM5Zf	Gene replacement constructs	Promega
pLEW100	Protein expression under tetracycline-inducible control	Wirtz *et al*
pLEW82	Protein expression under tetracycline-inducible control	Wirtz *et al*
pMOTag4YH	*In situ* tagging	Oberholzer *et al*
pMOTag4H	*In situ* tagging	Oberholzer *et al*
pSC-A-amp/kan	General cloning	StrataClone
pXS5	BipN generation	Jay Bangs

2.3.3 PCR

All polymerase chain reactions (PCRs) for amplification of UTRs, ORFs, and Southern probe were performed using a PTC-225 Peltier Thermal Cycler (MJ Research). Reactions were run with the Kod Hot Start Polymerase in a 50 µl volume as per the manufacturer's standard protocol using the appropriate primers (Table 2.1). The template concentrations used ranged from 1-10 ng for gDNA and 10-100 pg for purified plasmid DNA and optimised for each target of interest.

When joining the UTRs in the gene replacement construct, the two UTRs were first annealed via the previously introduced complementary sequence (in the absence of primers) followed by amplification of the combined UTR fragment by the addition of the 5' UTR Fw and 3' UTR Rev primers (Table 2.5).

Table 2.5. Annealing PCR reaction for joining UTRs

Reaction Mix (per 50 µl)
5 µl of 10x KOD Hot PCR Buffer
5 µl dNTPs (0.2mM each)
2 µl $MgSO_4$(1mM)
50 ng each UTR fragment
1 µl KOD Hot Start DNA polymerase (1U)
up to 50 µl with nuclease-free H_2O
after linking PCR add:
1.5 µl 5' UTR – F primer (15 pmol)
1.5µl 3' UTR – R primer (15 pmol)

Amplification Program			
	#of Cycles	T (ºC)	Time
	1	95	30 sec
x5	1	95	30 sec
	1	50	45 sec
	1	68	1 min
	1	95	30 sec
x25	1	95	30 sec
	1	60	45 sec
	1	68	1 min
	1	68	10 min

2.3.4 Agarose gel electrophoresis

Agarose gel eletrophoresis was used for separation of DNA. Gels for general DNA analysis were made with 1 % w/v agarose (Electran agarose, VWR) in 1x TAE (40 mM Trisacetate, 1 mM EDTA) buffer containing 0.4 μg/ml ethidium bromide, while 0.8% gels were used for Southern blotting. The gels were run at 80 V in 1x TAE, with the exception of agarose gels for Southern blotting that were run at 40 V,until the desired separation was achieved. The separation of a DNA marker (1 kbp DNA ladder, Promega) allowed the size of the samples to be estimated by comparing the relative migrations through the gel. Gels were imaged by UV transillumination.

2.3.5 Gel extraction of DNA fragments and plasmid DNA

To purify DNA fragments and plasmid vectors obtained from PCR or endonuclease digestions following separation on agarose gel, QIAquick® Gel Extraction Kit (Qiagen) was used following kit intructions. DNA concentration was estimated by A260 at the spectrophotometer (Nanovue).

2.3.6 Transformation

Transformation of plasmids was achieved by heat shock of competent cells. Typically, 10 ng of purified plasmid DNA was mixed with 50 ul chemically competent cells (Bioline α-select silver, Agilent technologies XL10 gold, StrataClone Solopack) and incubated on ice for 30 minutes before being subjected to heat chock at 42°C for 30

seconds followed by cooling the cells on ice for 2 minutes. Subsequently, SOC medium was added to the cells which were then grown 1 h at 37 °C, 200 rpm before being plated on LB agar plates containing ampicillin. Colonies were selected after overnight growth at 37 °C.

2.3.7 *E. coli* strains and cultures

E. coli cells (DHα5, Top10 one shot, Silverselect – Invitrogen and Novagen clonable cells) were grown in LB medium containing 50 μg/ml ampicillin at 37°C, 200 rpm. After transformation, colonies were selected on LB agar plates grown at 37°C for 16 h. For bacterial stabilates, cells were mixed 1:1 with 80% glycerol (AnalaR, VWR) and freezed on dry-ice before storage at -80°C. All media were obtained from the Media Kitchen service in theCollege of Life Sciences, University of Dundee.

2.3.8 Dephosphorylation

Following endonuclease digestion, plasmids were dephosphorylated to reduce the re-ligation. The plasmids were incubated at 37°C for 30 min with Antartic phophatase (New England Biolab) following by inhibitionof the reaction by heating at 65°C for 20 min.

2.3.9 Ligation

Ligations were made by using the Rapid DNA Ligation kit (Roche) or the Clonables kit(Merck Millipore). Typically, the concentrations of digested vector and inserts were estimated at the spectrophotometer (A260) and by running an analytical agarose gel

where the DNA bands could be compared. Vector and inserts were mixed accordingly to the manufacture's protocols and an additional reaction without insert was run as a control for the vector re-ligation background.

2.3.10 DNA sequencing

DNA sequencing was performed by the Sequencing Service, University of Dundee. The consensus sequence was built using or CLC Main Workbench 6.9.1.

2.3.11 Site-Directed Mutagenesis

Site directed mutagenesis was used to introduce a TEV site into the pMOTag4YH and pMOTag4H plasmids in addition to generate point mutations at position 397 in *Tb*STT3B and *Tb*STT3C. The QuickChange® Site-Directed Mutagenesis Kit (Stratagene) was used according to manufacturer's instructions. The primers used are listed in Table 2.1.

2.3.12 Ethanol precipitation of DNA

Ethanol precipitation of genomic DNA was done using 3 M sodium acetate pH 5 (1/10 of DNA solution volume) which was supplemented by absolute ethanol (3 times the total volume of DNA plus sodium acetate). The mix was incubated at -20°C for at least 16 h or alternatively 30 min on dry ice. To resuspend the precipitated DNA, the mix was centrifuged at 4 °C for 10 min at 16000 g and the supernatant was decanted followed by the pellet being washed with ice cold 70% ethanol. After decanting the wash, the ethanol was left to air-dry and then the pellet was resuspended in buffer or

sterile water. Finally, purified samples were confirmed to be pure by spectrophotometry (260:280 ratio >1.8).

2.3.13 Genomic DNA isolation

Bloodstream form *T. brucei* were grown to log phase density and ~1x10^8 cells were harvested by centrifugation for 10 min at 4 °C, 800 g. Cells were then lysed in 1 ml DNAzol™ and incubated at least 30 min at room temperature. Subsequently, 0.5 ml of absolute ethanol was added and the cell lysate was centrifuged for 10 minutes at 16000 g. The supernatant was decanted and the remaining pellet washed twice with 70% ice cold ethanol (10 min, 16000 g). Subsequent to decanting the wash, the pelleted genomic DNA was left to air-dry (~10 min) before being dissolved in 50 µl of H_2O and heated to 50 °C for about 4 h.

2.3.14 Southern blotting

Genomic DNA (~ 5 ug) was digested for 16 h with specific restriction endonucleases and run (40 V) on a 0.8 % agarose gel containing ethidium bromide. Afterwards, the gel was washed 5 minutes in 1x TAE before continuing washing it 10 minutes in 0.2 M HCl, 15 min in 0.5 M NaOH, 1.5 M NaCl and twice for 20 min in 1 M TrisHCl, 1.5 M NaCl pH 7.5 to prepare for transfer. The DNA was transferred to a positively charged nylon membrane (Roche) by capillary mechanism as described in Figure 2.1. The transfer was performed overnight in 10x sodium chloride, sodium citrate buffer (SSC). Subsequently, the DNA was crosslinked to the membrane using 1200 UV counts on a CL-1000 (UVP) UV crosslinker.

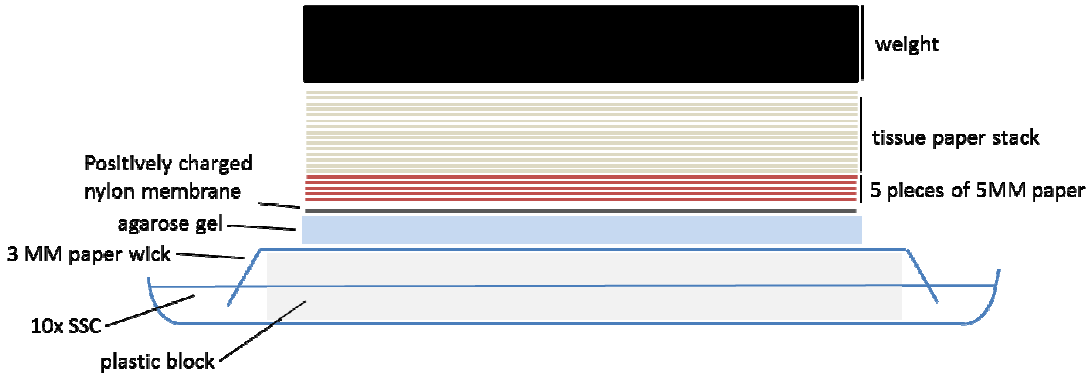


Figure 2.1. Southern blot transfer set up of fragmented gDNA from the agarose gel to the nylon membrane.

2.3.15 DIG-labelled probe detection

The DNA probes used were generated using the PCR DIG Probe Synthesis Kit (Roche). A 5 ul aliquot was analyzed by agarose gel electrophoresis to control that a single band was produced and that there was a size increase in the labelled probe compared to unlabelled negative control. When a bright single band was observed, 45 µl of probe /45 ml of hybridization buffer were used. The membrane was placed in a roller bottle (Techne) and pre-hybridized in EasyHyb solution (Roche) at 42°C for 30 min in a Hybridized HB-1D oven (Techne). Before use, the probe was denatured by heating for 100°C for 5 min and quickly cooled on ice, and subsequently added to a fresh aliquot of EasyHyb solution. The membrane was then hybridized overnight at 42°C. Following hybridization, the membrane was washed twice for 5 min in 1x SSC, 0.1% SDS at 42°C and twice for 15 min in pre-warmed 0.5x SSC, 0.1% SDS at 65°C. The membrane was then equilibrated for 5 min at room temperature in 1x wash buffer before being blocked for 30 min in blocking solution (DIG wash and block buffer set, Roche). After blocking, the membrane was incubated for 30 min with Anti-DIG AP-conjugate antibody

(Roche)at a 1:10000 dilution in blocking buffer of, followed by two 15 min washing steps in 1x wash buffer and a 5 min equilibration step in 1x detection buffer. The membrane was placed in a plastic folder and CSPD detection reagent (Roche) was applied in a drop wise manner. The membrane was incubated 5 min at room temperature in the dark and 10 min at 37°C. Films (Hyperfilm™ ECL, Amersham) were exposed from 30 sec to 5 min and developed with a Compact X4 developer.

2.4 Protein biochemistry

2.4.1 SDS-PAGE

All SDS-PAGE were run using pre-cast Novex Bis-Tris gels (4-12%, 10% or 12%) with MOPS running buffer (50 mM MOPS, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.7 - Invitrogen) in n XCELL-II mini tank (Novex) at 200 V using a BioRad Power Pac Junior. Samples were prepared using 4x SDS Sample buffer (200mM Tris-HCl pH 6.8, 400 mM DTT, 8 % SDS, 0.4 % bromophenol blue and 40 % glycerol – Invitrogen) and a final concentration of 0.1 M DTT as reducing agent. Before loading, samples were heated at 100°C for 10 min.

2.4.2 Coomassie staining

After desired separation was achieved, proteins were stained by incubating for 1 h with 0.1% Coomassie brilliant blue stain in 40% methanol, 10% acetic acid. The excess stain was removed by washing the gel in destain solution (40% methanol, 10% acetic acid) or in H_2O. For peptide mass fingerprinting, the gel was stained for 1 h min with SimplyBlue™ Safe Stain (Invitrogen) and destained in water.

2.4.3 Western blotting

After separation by SDS-PAGE, proteins were transferred on a nitrocellulose membrane (iBlot gel transfer stack) using a iBlot system (Invitrogen) according to manufacturer's instructions. The membrane was incubated for 5 min in Ponceau S (Sigma) before being blocked in 50 mM TrisHCl, 0.15 M NaCl, 0.05 % Tween 20, 0.25 % BSA,0.05 % Na and 2 % FSG pH 7.4 (Blocking buffer) for 20 min. The membrane was then incubated for 30 min with primary antibody in a 50 ml Falcon tube followed by washing using the SnapID system. Subsequently, the secondary antibody was incubated for 30 min and again followed by a washing step. Both these steps were carried out using the SnapID.

2.4.4 Western blot development

Proteins separated on SDS-PAGE gels were imaged using an ODYSSEY® SA near infrared imager (LI-COR Biosciences).Secondary LI-COR antibodies (IRDye-800CW goat anti-mouse 1:15000 or IRDye-680RD donkey anti mouse 1:20000) were used to bind the primary antibodies that had bound to the specific tags. Gels were scanned with a resolution of 50 μm using the appropriate sensitivity settings for each channel and a focal path length of 3.5 mm. Coloured images were sometimes converted to grayscale and, when needed, proteins were quantified in Image Studio Lite version 3.1 (LI-COR Biosciences).

2.4.5 Immunoprecipitaion

100 ml of cells culture were grown to log phase (approximately $2.5x10^6$ cells/ml) and lysed in lysate buffer (see table 2.7) for 30 min on ice (Lysis buffer 2 and 3) or by boiling for 10 min (lysis buffer 1). Subsequently, the lysate was centrifuged (4°C, 11000g, 15 min) and the supernatant was moved to a new tube. When using lysis buffer 1, the lysate was diluted 1:30 in 50 mM Tris HCl pH 6.8, 20 mM EDTA and the protease inhibitors mentioned earlier. Pre-washed protein G magnetic beads were added to the lysate in order to absorb a majority of the nonspecific binding to the beads. After 30 min incubation (4°C with slow spin) the magnetic beads were captured and the lysate moved to a new tube followed by an hour incubation with antibody (4°C, slow spin). The different antibodies used are listed in table 2.6. Next, the magnetic beads were again captured with a magnet and the flow through removed before washing twice with 1 ml lysis/IP buffer. The final wash was done in wash buffer 2 (10 mM Tris-HCL, 4 mM EDTA, 0.1 % digitonin) containing the same protease inhibitors as mentioned earlier. To elute, 15 ul 1xPBS buffer, 4x sample load and 1 ul 1 M DTT was added to the beads, followed by boiling at 100°C for 10 min. The eluted proteins where subsequently separated on a SDS-PAGE gel.

Table 2.6. List of antibodies used for immunoprecipitation.

Antibody	Concentration	Supplier
αHA	1µg/ml	In house
αHA precoupled	1µg/ml	Pierce
αMYC precoupled	1µg/ml	Pierce
αGFP	1µg/ml	Roche

Table 2.7. List of lysis buffers.

Lysis buffer 1	1 % SDS, 0.1 M DTT, PMSF, TLCK, Mini EDTA free protease inhibitor tablet
Lysis buffer 2	1 % TritonX, 50 mM Tris HCl pH 6.8, 20 mM EDTA, PMSF, TLCK, Mini EDTA free protease inhibitor tablet
Lysis buffer 3	0.5 % digitonin, 50 mM Tris HCl pH 6.8, 20 mM EDTA, PMSF, TLCK, Mini EDTA free protease inhibitor tablet

2.4.6 Blue native gel and Western blotting

Blue Native gel electrophoresis was run using components from the Native Page kit (Invitrogen). The protocol was followed to the manufacture's instructions with a few exceptions. No G-250 sample additive was added to the protein samples. The 1x NativePage Light Cathode buffer was diluted 1:4 in 1x running buffer to reduce coomassie interference for the following Western blot and image development.

2.5 SILAC

2.5.1 SILAC labelling

SILAC IMDM media (Thermo Scientific) was supplemented with the ingredients listed in table 2.8. After adjusting the pH to 7.4, either normal L-arginine and L-lysine (HMI11-SILAC + R_0K_0) or L-arginine U-$^{13}C_6$ and L-lysine 4,4,5,5-$^{2}H_4$ (HMI11-SILAC + R_6K_4) was added to the media (120 and 240 uM respectively) which was then sterile filtered (TPP vakuumfiltration 500). Cells were split into either heavy or light labelled SILAC media with a dilution factor of 1:1000 and cells were subsequently grown to log phase before starting the IP experiment.

2.5.2 SILAC IP

Heavy and light label cells were spun down individually (15 min, 800 g, 4 °C), washed in TDB before being resuspended in 1 ml 1xTDB for accurate cell counting. The cells were then mixed 1:1 before treated as described in the IP protocol, using lysing buffer 1. When to elute, 12 µl wash buffer 2 (10 mM Tris-HCL, 4 mM EDTA, 0.1 % digitonin), 10 µl 4x loading buffer and 5 µl 50 mM DTT (Clealand reagent) was added to the captured beads and the mixture was boiled at 100°C. 2 µl of the sample was analysed by Western blotting to verify that the tagged protein had been immunoprecipitated and, if so, the remaining 25 µl was handed to the proteomics facilities for further preparations.

2.6 Mass spectrometry and data processing

2.6.1 Mass spectrometry

All trypsin digestion and mass spectrometry was carried by the FingerPrints proteomic service at the University of Dundee (http://proteomics.lifesci.dundee.ac.uk)
SILAC protein samples were alkylated by adding 5 µl of 300 mM iodoacetamide and incubation at room temperature for 30min in the dark. Samples were loaded on novex NUPAGE 4-12% Bis Tris gel and run at 200 V using MOPS buffer until the protein have migrated 2 cm into the gel. Proteins were stained with simply blue safe stain for 1hr and then destained by gentle shaking in MilliQ water followed by in gel digestion. Protein bands were cut with a clean scalpel blade, transferred to a micro-centrifuge tube and cut into small pieces. Gel pieces were washed with Milli-Q water (200 µl), 200 µl of acetonitrile, 200 µl of 100 mM ammonium bicarbonate, 200 µl 100 mM ammonium bicarbonate and 50% acetonitrile and finally 100 µl acetonitrile for 15 min each step.

Gel pieces were then air dried in a laminar flow cabinet. In-gel digestion is carried out by 1µg (1µl) of trypsin, incubated overnight at 30^{o}C. An equal volume of acetonitrile (i.e. 100 µl) was added and the peptide mixtures were extracted by vigorous shaking (1,000 rpm) at room temperature for 10 min. The resulting supernatant was then transferred to a fresh microcentrifuge tube. Peptides in the gel pieces were further extracted twice with 5% formic acid (100 µl) and 100% acetonitrile (100 µl) with 20 min incubation each on the shaker. The supernatant was then collected and transferred to the first fraction. The gel pieces were washed for a third time with acetonitrile (100 µl) and incubated for a further 10 min. The pooled three fractions were dried in SpeedVac and stored at -20^{o}C until further processing.

Following digestion, peptides from various samples were reconstituted in 50 µl 1% formic acid and 10 µl were analysed on a Velos pro-orbitrap (Thermo Scientific) mass spectrometer coupled with a Dionex Ultimate 3000 RS (Thermo Scientific). LC buffers were the following: buffer A (2% acetonitrile and 0.1% formic acid in Milli-Q water (v/v)) and buffer B (80% acetonitrile and 0.08% formic acid in Milli-Q water (v/v). Aliquots of 2 µL of each sample were loaded at 5 µL/min onto a trap column (100 µm × 2 cm, PepMap nanoViper C18 column, 5 µm, 100 Å, Thermo Scientific) equilibrated in 98% buffer A. The trap column was washed for 3 min at the same flow rate and then the trap column was switched in-line with a Thermo Scientific, resolving C18 column (75 µm × 50 cm, PepMap RSLC C18 column, 2 µm, 100 Å). The peptides were eluted from the column at a constant flow rate of 300 nl/min with a linear gradient from 98% buffer A to 40% buffer B in 128 min, and then to 98% buffer B by 130 min. The column was then washed with 98% buffer B for 20 min and re-equilibrated in 98% buffer A for 29 min. LTQ-Orbitap Velos was used in data dependent mode. A scan cycle comprised MS1 scan (m/z range from 335-1800) in the velos orbitrap followed by 15 sequential dependant MS2 scans (the threshold value was set at 5000 and the minimum injection

time was set at 200 ms) in LTQ with collision induced dissociation. The resolution of the Orbitrap Velos was set at to 60,000. To ensure mass accuracy, the mass spectrometer was calibrated on the first day that the runs were performed.

2.6.2 Tryptic mass fingerprinting

Coomassie stained SDS-PAGE gels were digested in-gel with Trypsin gold (Promega) and analysed by LC-MS/MS. In-solution digestion of immobilised proteins was also performed with trypsin gold and the peptides recovered and analysed by LC-MS/MS. LC-MS/MS spectra were acquired on an LTQ Orbitrap XL (Thermo Fisher).

2.6.3 MaxQuant

Data were processed using MaxQuant version 1.3.0.5 which incorporates the Andromeda search engine (Cox and Mann, 2008). Proteins were identified by searching a protein sequence database containing *T. brucei brucei* 927 annotated proteins (Version 7.1, downloaded from TriTrypDB http://www.tritrypdb.org) supplemented with the VSG221 sequence and those offrequently observed contaminants (porcine trypsin, bovineserum albumin, and human keratins). Search parametersspecified an MS tolerance of 10 ppm, an MS/MS toleranceat 0.6 Da, and full trypsin specificity, allowing for up to twomissed cleavages. Carbamidomethylation of cysteine was set asafixed modification and oxidation of methionines, N-terminalprotein acetylation, andN-pyroglutamate were allowed asvariable modifications. SILAC ratioswere calculated using only peptides that could be uniquelymapped to a given protein group and required a minimum of 2SILAC pairs. The distribution of SILAC was

normalized withinMaxQuant at the peptide level so that the median of log_2Ratiois zero to account for any errors in cell counting.

2.7 Microscopy

2.7.1 Cell fixing, staining and imaging

All solutions used in this protocol were filtered on a 0.2 μm Minisart filtering unit (Sartorious). About 2.5x10^7*T. brucei* bloodstream form cells at log phase density were harvested by centrifugation and washed with cold 1x TDB followed at 600xg, 10 min at 4°C (4K15 Sigma Centrifuge) and resuspended very gently to reduce cell bursting.
Gently resuspend in 1ml cold PBS and add an equal volume of 8% PFA at 4°C for 30min. Place 10 µl of cell culture on the centre of pre-washed coverslip and let air-dry in hood at RT. Coverslips was pre-washed in 70% EtOH and air-dried. Wash in PBS+ 0.05% NaN_3 and keep wet at 4°C until needed. To permeabilize, incubate the coverslips 10 min at RT, shaking in 1 ml of filtered 0.1% TX-100 in PBS (inside the 12-well plate). Wash the coverslips once with 1 ml of 1x PBS (filtered) and transfer the coverslips on top of parafilm inside a humid chamber. Block 1h at RT with 50µL of 5% (w/v) Fish Skin Gelatin (FSG, Sigma G7765-250ml) + 0.05% (w/v) TX-100 + 10% normal goat serum in PBS containing 0.05% NaN_3 (pre-filtered in 0.2 um). Pipette carefully up & down, so that the solution is well distributed. Spin the goat serum solution to remove any precipitate before adding to the blocking solution. The goat serum was used because the Alexa secondary antibodies have been produced in goat and adding it to the blocking buffer should help reduce unspecific signal. Wash 1x 50ul of filtered 1% w/v FSG + 0.05% (w/v) TX-100 in PBS containing 0.05% NaN_3. Prepare the 1st antibody solution by diluting the antibody stock in filtered 1% (w/v) FSG +

0.05% (w/v) TX-100 in PBS containing 0.05% NaN_3. If using more than one primary antibody, dilute them together in the same solution. After the dilution, spin 5 min at 16,000 xg, 4˚C to remove insoluble materials. Add 50µl/ coverslip of antibody solution and pipette very carefully up & down to be sure that the solution is well distributed. Incubate 1 hour at RT. Then wash 3x 50µl of filtered 1% w/v FSG + 0.05% TX-100 in PBS containing 0.05% NaN_3. Prepare and apply 2nd antibody solution as described for the 1st antibody and cover the humid chamber with foil during the incubation (keeping light away). Subsequently, wash 3x 50 µl of filtered 1% w/v FSG + 0.05% TX-100 in PBS containing 0.05% NaN_3. Wash with 2x 50ul of filtered TBS, remove it by aspiration. For mounting, take a frosted glass slide and place 4-5µl of Prolong Gold onto the glass slide, then aspirate the last wash from the coverslip, and immediately place the side containing the cells down, in contact with Prolong Gold. Incubate the slides 5-10 min in the dark. Seal the coverslips with nail varnish and store them covered in foil in the cold room, until microscope availability

2.8 Sequon specificity studies

2.8.1 *Tb*BipN reporter protein

BipN was PCR amplified by using primers previously published in (Bangs et al., 1996) (seen intable 2.1) and cloned into a pLEW82 plasmid. A glycosylation sequon sequence with a HA_3 tag (ordered from Dundee Cell Products, seen in appendix) was subsequently cloned in directly downstream the BipN sequence and the construct was transfected into wild type *Trypanosoma brucei* cells.

2.8.2 *Tb*BipN glycan digestion

10 ml of mid log phase (~$2x10^6$) cells with overexpressed BipN reporter protein were lysed using 40 µl of lysing buffer 1 (section 2.4.5) and heated at 100°C for 10 min before being cooled to room temperature and diluted 1:10 in PBS buffer. To 5 ul of sample, 7.5 µl of 1x G5 buffer (NEB) and 7.5 µl (37.5 mU) of EndoH was added over night for digestion. PNGase F is compatible with the PBS buffer and therefore 5 µl sample was mixed with 7.5 µl enzyme (7.5 U) and 7.5 µl H_2O for digestion over night.

2.8.3 *Tb*BipN glycosylation sequon specificity

7 µl of 4x loading buffer and 0.5 µl of 1 M DTT respectively 10 % SDS was added to the digested samples before heating at 95 °C. Samples were loaded on a 12 % pre-casted NuPage gel and run until the 37 kDa marker (All blue, Bio Rad) ran off the gel, in order to separate the small changes of molecular weight of the reporter protein. Subsequently, the gels were transferred to nitrocellulose membrane as described in section 2.4.3 and the Western blot was developed according to section 2.4.4. Finally, the bands were quantified using Image Studio Lite version 3.1 (LI-COR Biosciences).

2.8.4 *T. brucei* STT3 models

Models were built using Phyre2 software. The amino acid sequence of TBsTT3A, TbSTT3B and TbSTT3C were inserted and the top hit was the crystal structure of *C. lari* (Lizak et al., 2011). The models were visualised in PYMOL.

2.8.5 *Tb*STT3A and *Tb*STT3B motif logos

The *Tb*STT3A and *Tb*STT3B motif logos were generated using Weblogo3. Acceptor peptides that exclusively received $Man_9GlcNAc_2$ or $Man_5GlcNAc_2$ are seen in Appendix.

3 – Results

3.1 Essentiality of *Tb*STTA - generation of *Tb*STT3A gene replacement mutants

Previous RNAi gene knockdown showed that individual *Tb*STT3Aor *Tb*STT3B knockdowns (by 80-90%) produced dramatic glycosylation changes but did not affect cell viability in culture, whereas simultaneous knockdown of both these genes caused cell death (Izquierdo et al., 2009a). Sincethe RNAi knockdown was incomplete,there are still questions as to whether *Tb*STT3A and *Tb*STT3B might be individually essential for cell growth *in vitro*. Therefore, to continue the work of Izquierdo *et al*, and seek further knowledge concerning the essentiality of *Tb*STT3A *in vitro*, gene replacement experiments were initiated. The Lister 427 bloodstream form cell line used for all genetic modification has been genetically modified to expresses a T7 polymerase and a tetracyclin repressor protein (TetR) under neomycin control. These modifications make it possible to generate both null and conditional mutations by gene replacement (Wirtz et al., 1999).

3.1.1 Introduction of a *TbSTT3A* ectopic copy

Using the *Tb*STT3A,B,C+/- heterozygote cell line previously established (Izquierdo et al., 2009a), a gene replacement construct was engineered to try to specifically remove the last copy of *TbSTT3A* (data not shown). However, no surviving cells could be obtained after transfection and selection, so a new strategy for targeted replacement of *TbSTT3A*with an antibiotic cassette was designed. We reasoned that, if *Tb*STT3A is essential also *in vitro* (i.e. in cell culture),an ectopic gene copy expressing sufficient levels of the enzyme will be needed to allow replacement of the last endogenous

TbSTT3A gene. Therefore, a pLEW100 plasmid containing the *TbSTT3A* gene (previously engineered by Luis Izquierdo) was transfected into the heterozygote cell line. This was the first step in trying to create a conditional null mutant. The pLEW100 plasmid is a dual promoter expression vector which targets the rRNA spacer locus (see Figure 3.1.2). The plasmid has a procyclin promotor which is driving expression of the *TbSTT3A* gene through two (for greater efficiency) tetracycline operator sequences. The pLEW100 plasmid also contains a T7 promoter which constitutively expresses the phleomycin resistance gene (Wirtz et al., 1999). Hence, phleomycin can be used for the selection of transfected cells after electroporation. The ectopic copy of *TbSTT3A* was not tagged.

Subsequent to obtaining recovered cells, the parasites were cloned and genomic DNA was extracted to test for the introduction of the *TbSTT3A* ectopic copy. With the three *STT3* genes sharing a large percentage of their DNA sequence and also being surrounded by identical untranslated region (UTR) repeats, Southern blotting proved to be challenging. However, by probing against a stretch of sequence in *TbSTT3A*(primers found in table 2.1) it was possible to verify introduction of the ectopic *TbSTT3A* copy. Genomic DNA from transfected (5 µg) and parental (2 µg) cell lines were digested with PstI and BglII. PstI is found upstream of the *STT3* locus whereas BglII is found in the 3' end of all three genes. In the pLEW100 plasmid a BglII site is found upstream of *TbSTT3A*as well as in the C-terminus of the gene.(Figure 3.1.1).

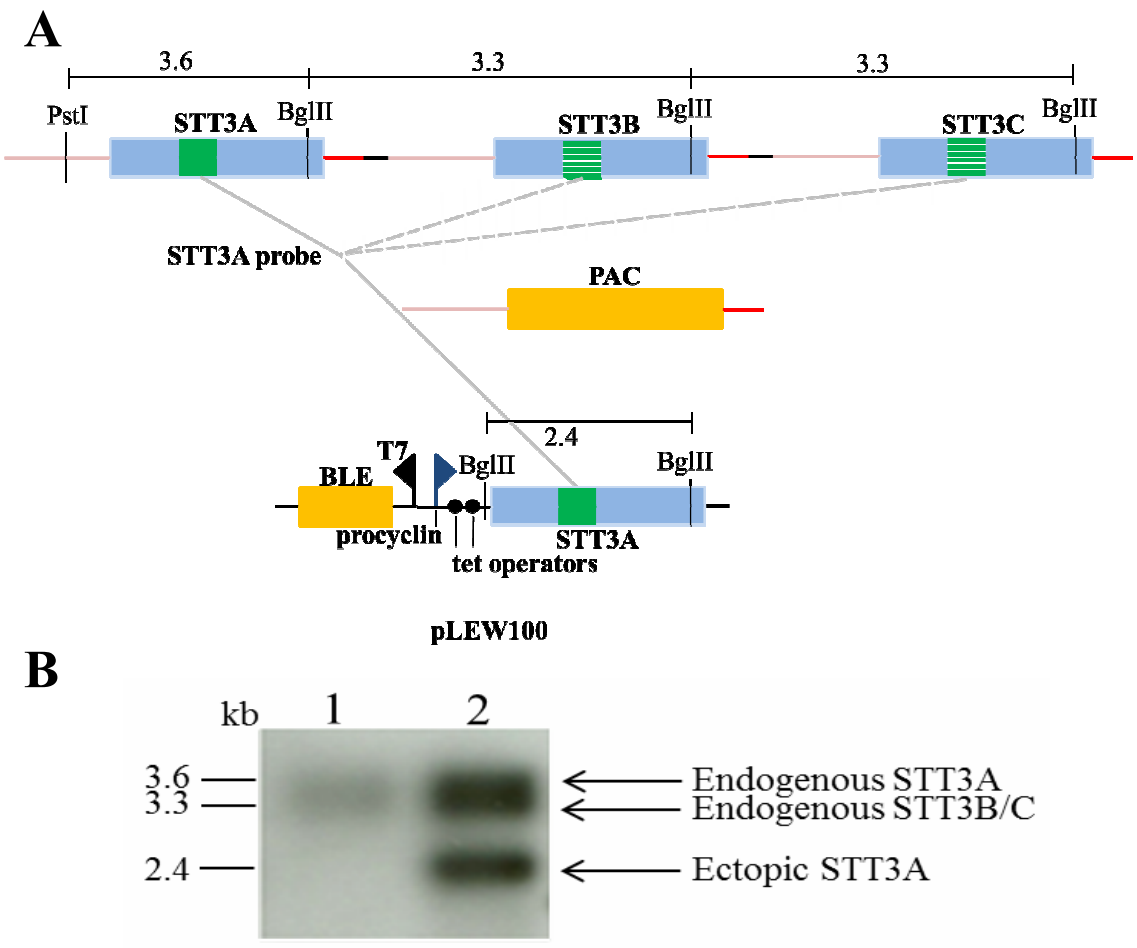


Figure 3.1.1. A. Digestion pattern for Southern blot analysis. A schematic showing that the probe binds best to a sequence in *TbSTT3A* (green box) and gives rise to bands at 3.6 kb (endogenous *TbSTT3A*) and 2.4 kb (ectopic *TbSTT3A*). The probe still associates weakly with *TbSTT3B* and *TbSTT3C*(green and white striped box), providing the 3.3 kbband. **B.** Southern blot showing the endogenous *Tb*STT3A,B and C genes in the parental *Tb*STT3A,B,C +/- cell line (lane 1) and the introduction of the ectopic copy of *Tb*STT3A in the *Tb*STT3A,B,C +/-; *Tb*STT3A tetracycline-inducible cell line.

The single additional band was of the correct size and confirmed that the ectopic *TbSTT3A* gene had been introduced to the rRNA spacer, establishing the cell line illustrated in Figure 3.1.2.

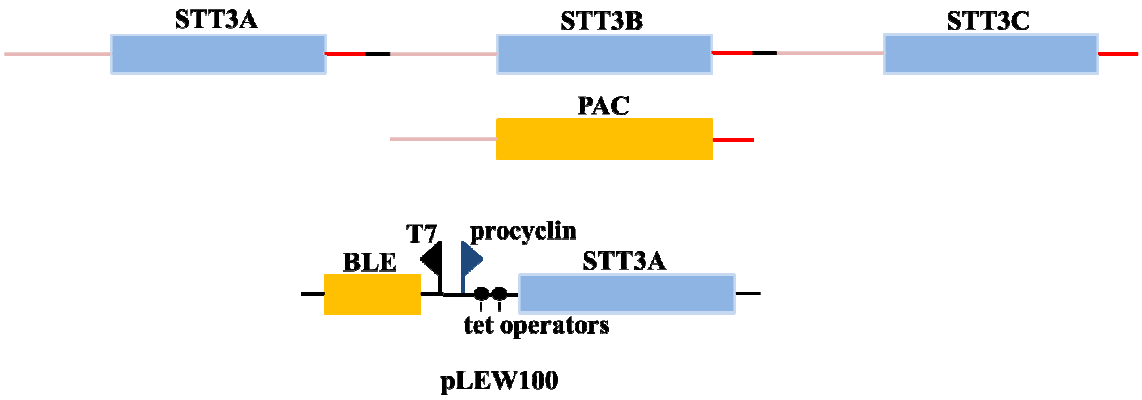


Figure 3.1.2. Cell line established by introduction of a tetracycline-inducible ectopic copy of *TbSTT3A*.

3.1.2 *In situ* tagging *TbSTT3A* with C-terminal YFP-HA_3

Following introduction of the ectopic *TbSTT3A* gene, a strategy for simplifying specific gene replacement of *TbSTT3A* was designed. By *in situ* tagging the endogenous *TbSTT3A*, unique DNA sequence will be introduced into the otherwise identical UTR repeats. This unique sequence can be used to specifically replace the last copy of *TbSTT3A*. The *in situ*tag will also enable immunoprecipitation and immunofluorescence experiments with the protein (chapter 3.2). Since the first transmembrane region of *Tb*STT3A act as a signal and anchor sequence that both targets the protein to the ER and anchor it to the bilayer, tagging was directed to the C-terminus which points into the ER lumen (Kim et al., 2005). pMOTag plasmids are generally used to *in situ* tag protein in *T. brucei* and were therefore chosen for this purpose. The plasmid series includes different tags and selection markers that can be used for tagging endogenous

genes, rendering native expression levels of the tagged proteins (Oberholzer et al., 2006). The pMOTag4YH introduces an YFP tag followed by an HA_3 tag under hygromycin (HYG) selection and was used for *in situ* tagging *Tb*STT3A. The 3' end of *TbSTT3A* (without the stop codon) including the stretch of unique sequence was PCR amplified together with the 3' UTR directly downstream of the gene. Both DNA products were subsequently cloned into the plasmid (primers found in table 2.1). The plasmid was linearized by NotI before being transfected into the previously established cell line (Figure 3.1.3).

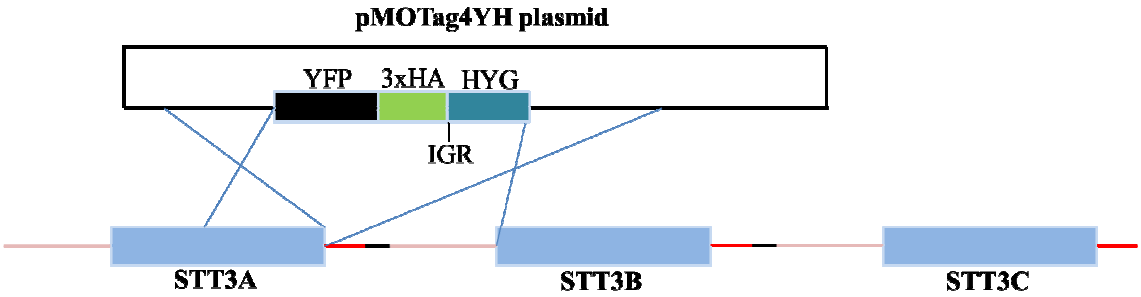


Figure 3.1.3. Strategy for the generation of the pMOTag4YH *in situ* tagging construct. The 3' end of *TbSTT3A* and the 3' UTR was PCR amplified and cloned into the pMOTaG4YH plasmid.

The unique sequence found in *TbSTT3A* should direct the tag to the 3' end of the gene (Figure 3.1.4). However, there was a risk that the 3' UTR sequence would homologously recombine downstream of *TbSTT3B*. In that case *TbSTT3A* would be *in situ* tagged simultaneously as the *TbSTT3B* gets knocked out.

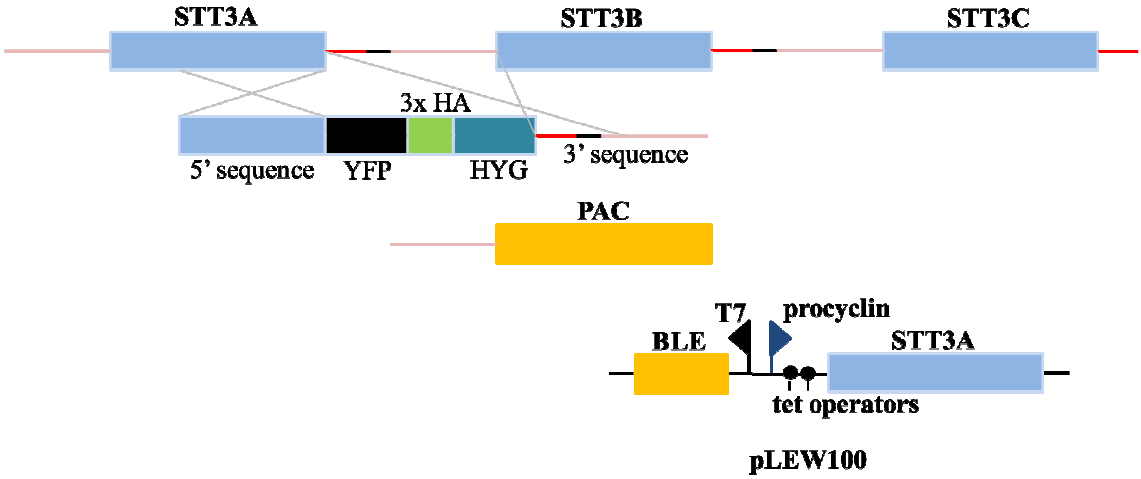


Figure 3.1.4. Strategy for the *in situ* tagging of *TbSTT3A*.The 3' end of *TbSTT3A* wasfused to an YFP tag followed by a HA_3 tag under HYG selection.

Since it was not known if the C-terminus tag would interfere with folding and/or function of the protein, the expression of the ectopic *Tb*STT3A was induced by supplementation of tetracycline to the media before transfection with the *in situ* construct. After HYG selection, recovered cells were made clonal and the genotype was analysed by Southern blotting. When probing against HYG following PstI or PstI/NdeI digestion (Figure 3.1.5), specific bands of the correct size were seen, confirming that *Tb*STT3A has been correctly *in situ* tagged. (Figure 3.1.6).

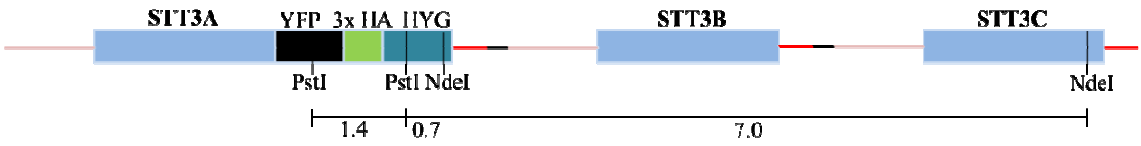


Figure 3.1.5. Digestion pattern for Southern blot analysis of *in situ* tagged *TbSTT3A*. When probing against HYG following PstI digestion, two bands should be observable at 1.4 respectively >10 kb. When digesting with PstI and NdeI, three bands should be seen at 0.7, 1.4 and 7.0 kb, respectively.

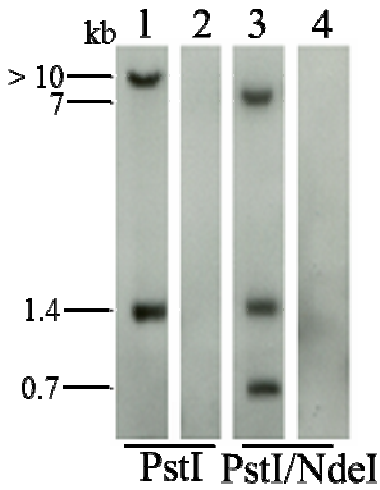


Figure 3.1.6. Southern blot visualising *in situ* tagging of *TbSTT3A*.Lane 1 and 3: Genomic DNA (5 µg) from transfected cells after PstI or PstI/NdeI digestion. Lane 2 and 4: Genomic DNA (5 µg) from parental cell line after PstI or PstI/NdeI digestion. The single PstI digestion resulted in two predicted bands, 1.4 and >10Kb, indicating that only one copy has been introduced. The PstI and NdeI double digestion released three HYG containing fragments of the predicted sizes 0.7, 1.4 and 7.0 kb. The 7Kb band confirmed that the *in situ* tag has been specifically introduced into the 3'-end of the *Tb*STT3A gene. A HYG probe was used for the analysis.

To further ensure that the *in situ* tag does not impair protein function of *Tb*STT3A, two additional analyses were carried out. By comparing the growth rate of cells in permissive and non-permissive conditions for eleven days it was evident that cells with the *in situ* tagged *TbSTT3A* gene as the sole source for the protein grew as well as the cells with the overexpressor induced (Figure 3.1.7).

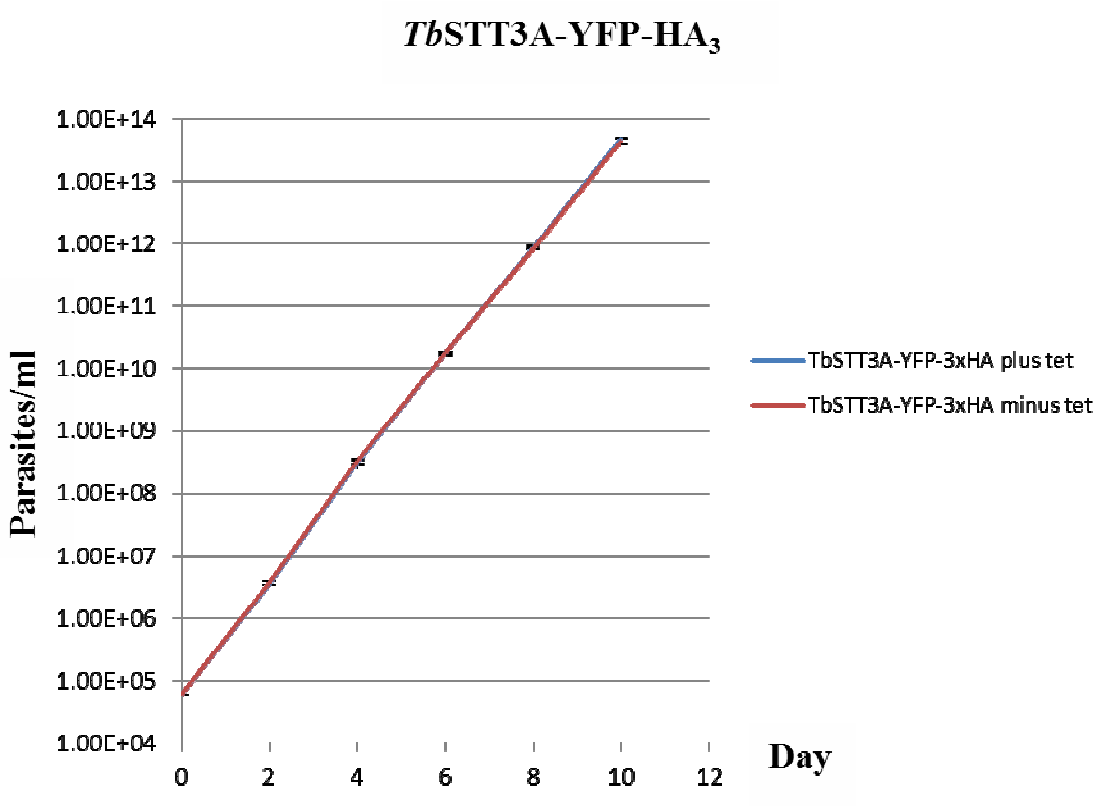


Figure 3.1.7. Growth curve of *in situ* tagged cells with or without overexpressed *Tb*STT3A from the ectopic copy. Cell numbers were counted in biological triplicates for eleven days and cells were split when required. Having *in situ* tagged *TbSTT3A* as the only source for the enzyme does not affect cell viability significantly.

However, since cells were viable during RNAi knock down of *Tb*STT3A, the function of the enzyme was also adressed. By purifying VSGs from cells in non-permissive conditions and subsequently digesting with PNGaseF and EndoH, conclusions about the glycosylation state could be drawn. The VSG found on parasites used for this thesis are of the 221 variant. It contains two N-glycosylation sites, one at Asn263 which is occupied by small, biantennary, endoglycosidase-H (EndoH)-resistant paucimannose and complex structures, which originate from the *Tb*STT3A transfer of $Man_5GlcNAc_2$ from the $Man_5GlcNAc_2$-PP-Dol precursor, and one at Asn428 occupied by conventional EndoH-sensitive triantennary oligomannose structures, which originate from the

*Tb*STT3B transfer of $Man_9GlcNAc_2$ from $Man_9GlcNAc_2$-PP-Dol(Izquierdo et al., 2009a; Jones et al., 2005; Zamze et al., 1991). Thus, analysis of VSG221 N-glycosylation allows us to determine whether the biantennary N-glycan is present, thereby verifying the function of the *in situ* tagged *Tb*STT3A. Cell-surface VSG was purified in a soluble form (sVSG221) by osmotic shock, a process that releases sVSG from the parasite surface by the action of endogenous GPI-specific phospholipase C which cleaves the dimyristoylglycerol lipid component of the VSG GPI anchor(Cross, 1975a, 1984; Ferguson et al., 1985). This cell surface-derived sVSG221 was subjected to both EndoH and PNGaseF digestion and analysed by SDS–PAGE and Coomassie blue staining. EndoH cleaves triantennary oligomannose oligosaccharides whereas PNGaseF digests oligomannose, hybrid, and complex oligosaccharides (Maley et al., 1989). These enzyme treatments will result in specific digestion patterns depending on the type(s) of glycans that have been transferred to the N-linked glycosylation sequons of the sVSG.In Figure 3.1.8, the same reduction in molecular weight is seen after EndoH and PNGaseF digestion as compared to a wild type cell line, showing that the *in situ* tagged *Tb*STT3A is still able to transfer $Man_5GlcNAc_2$ from the $Man_5GlcNAc_2$-PP-Dol precursor to the Asparagine at position 263 under non-permissive conditions.

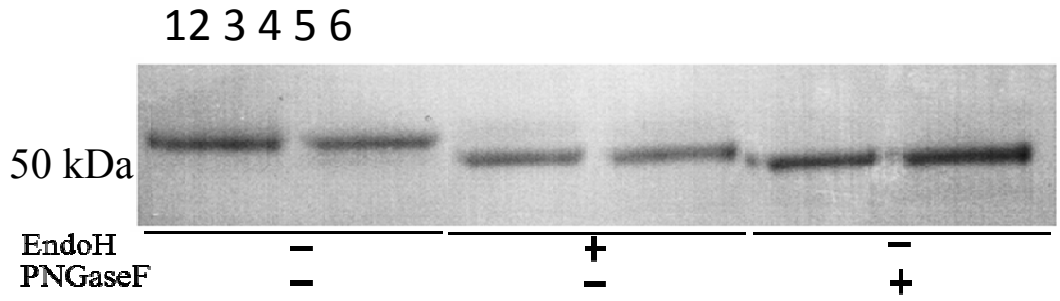


Figure 3.1.8.SDS-PAGE Coomassie stained gel visualising sVSG from the *Tb*STT3A *in situ* tagged cell line and a wild type control. EndoH cleaves off triantennary oligomannose glycans whereas PNGaseF digests both tri- and biantennary glycans. **Lane 1**: VSG221 from wild type cell line. **Lane 3 and 5**: VSG2212 from wild type cell line after EndoH (3) and PNGaseF (5) treatment. **Lane 2**: VSG221 from *in situ* tagged cell line. **Lane 4 and 6**: VSG221 from in situ tagged cell line after EndoH (4) and PNGaseF treatment.The digestion pattern indicates that $Man_5GlcNAc_2$ from the $Man_5GlcNAc_2$-PP-Dol precursor has been transferred to the asparagine at position 263 and, hence, that the function of the tagged *Tb*STT3A has remained unaffected under non-permissive conditions.

3.1.3 *TbSTT3A* gene replacement

By *in situ* tagging *Tb*STT3A, a unique DNA sequence was introduced to the *STT3* locus. This DNA was used in the attempts to replace *Tb*S*TT3A* with an antibiotic cassette under permissive conditions (Figure 3.1.9).

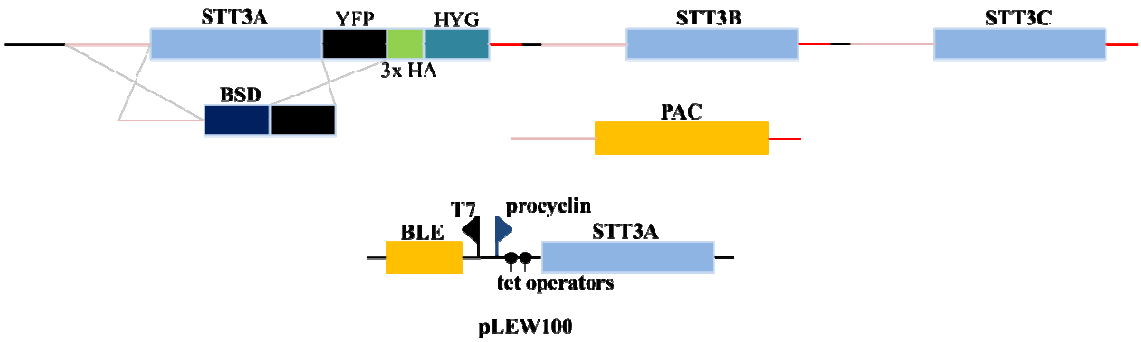


Figure 3.1.9. Strategy for *TbSTT3A* gene replacement under permissive conditions.

The approach to create the *TbSTT3A* replacement construct is described in Figure 3.1.10. Briefly, about 1000 bp of the 5' UTR and the flanking YFP were amplified by PCR (primers found seen in Table 2.1). The PCR products were joined together by PCR and cloned into the pGEM5Zf plasmid via two NotI restriction sites. Subsequently, the antibiotic resistance cassette was ligated into the plasmid via BamH and HindIII digestion. In a final step, the construct was released from the plasmid by NotI digestion and was later used in gene replacement trials.

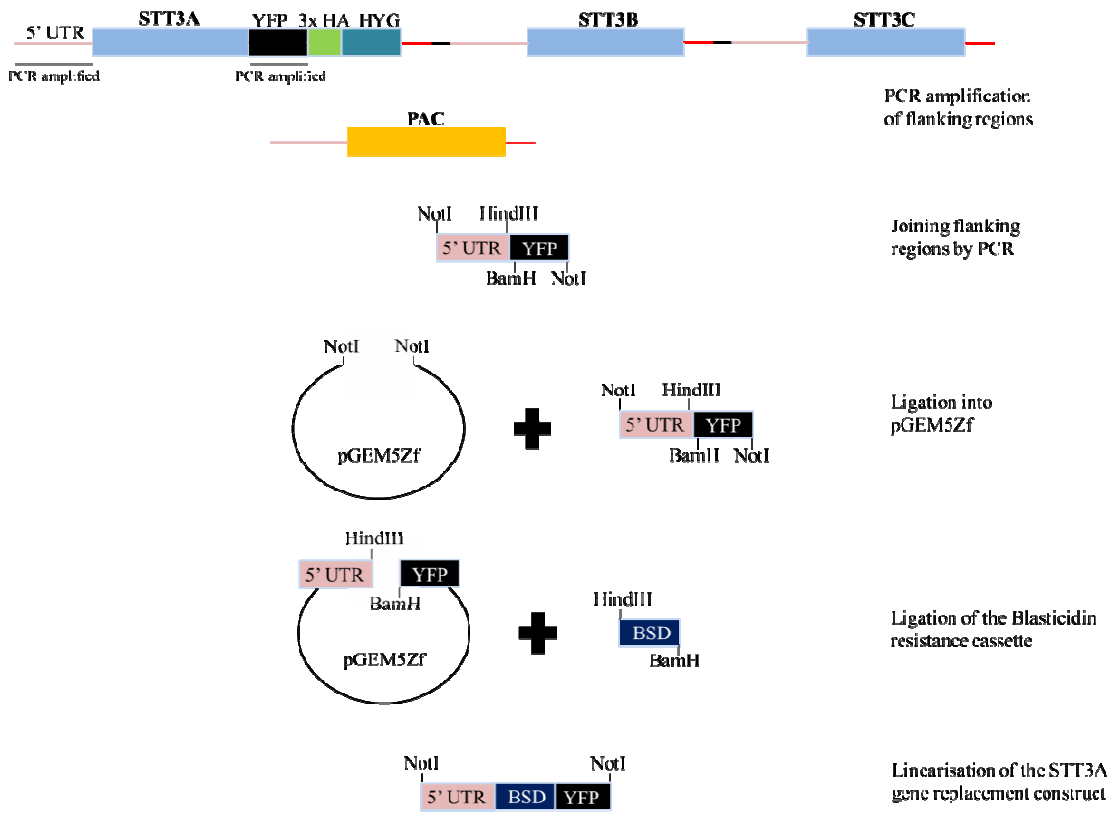


Figure 3.1.10. Strategy for the generation of *Tb*STT3A gene replacement construct.

After electroporation, recovered cells were made clonal and DNA was extracted for genomic analysis. A large number (20) of clones were initially screened by PCR. By using primers that anneal to a sequence outside the gene replacement construct, two different fragments could be amplified. If the 0.4 Kb BSD gene has replaced the 2.4 Kb *TbSTT3A* gene, a fragment of ~2 Kb should be observable, compared to a 4 Kb fragment if no replacement has occurred. Many (10) clones showed the smaller band, indicating that homologous recombination had occurred in the correct place (Figure 3.1.11).

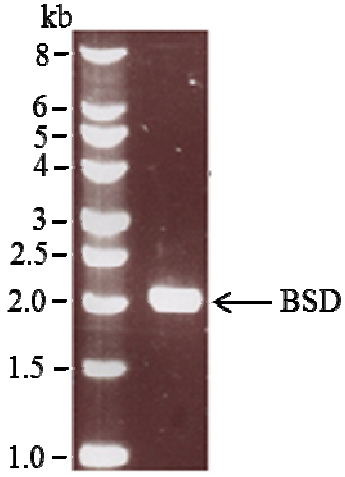


Figure 3.1.11.PCR screen of a *Tb*STT3A KO clone, showing replacement of *Tb*STT3A by BSD. Using primers that anneal outside of the gene replacement construct, two possible bands can be amplified. If *TbSTT3A* is still present, a 4 kb and will be amplified. If BSD has replaced *TbSTT3A*, a 2 kb band will be amplified.

Subsequent to the PCR screen, selected clones were followed up by Southern blot to verify the*TbSTT3A* knock out. In this experiment, a promiscuous *STT3* probe was used. This probe binds an identical sequence found in all three genes, *TbSTT3A, TbSTT3B* and*TbSTT3C*. By digesting with PstI, BspHI and XhoI four different sized fragments should be observable for the parental cell line. However, the 2.7 kb band should be absent in the clones if BSD has replaced *TbSTT3A* (Figure 3.1.12).

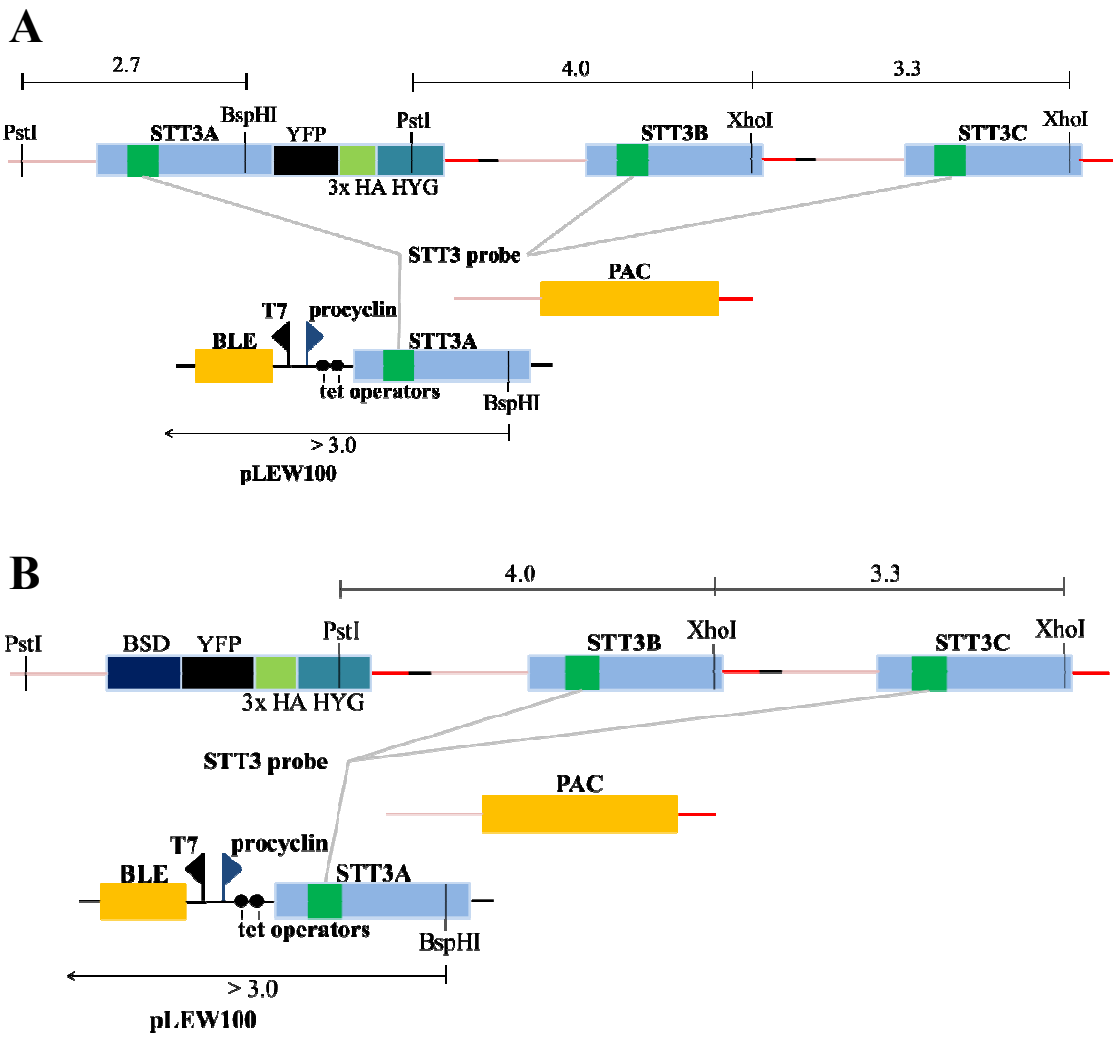


Figure 3.1.12. Digestion pattern for Southern blot visualising *TbSTT3A* gene replacement. The probe binds an identical sequence found in*TbSTT3A, TbSTT3B* and*TbSTT3C.* **A:** By digesting with PstI, BspHI and XhoI four fragments of different sizes should be visualised if *TbSTT3A* is still present). **B:** The 2.7 kb band should be absent if BSD has replaced *TbSTT3A*.

Figure 3.1.13 shows that the *TbSTT3A* band was lost in the knock out clones.However, a new band was seen, making interpretation difficult. Insufficient digestion was ruled out since the size of the new band could not be produced by the three restriction enzymes that were used. Since the ectopic *TbSTT3A* gene was not visible, questions were raised whether the previously introduced inducible ectopic copy of *TbSTT3A* had

been lost, possibly causing survival pressure to duplicate *TbSTT3A* in the genome. One way to address that question was to repeat the Southern blot with a different probe (*TbSTT3A* specific) and digestion pattern, providing information of fragment size of the ectopic *TbSTT3A* gene. However, the result showed that the inducible ectopic copy of *Tb*STT3A was still present (data not shown).

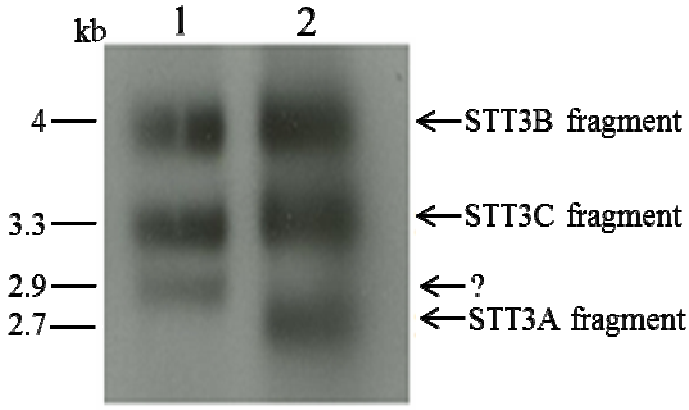


Figure 3.1.13. Southern blot visualising the loss of *TbSTT3A*.**Lane 1**: Genomic DNA (5 ug) from transfected cells after PstI, BspHI and XhoI digestion. **Lane 2**: Genomic DNA (5 ug) from parental cell line after PstI, BspHI and XhoI digestion. A promiscuous *STT3*probe binding an identical sequence found in*TbSTT3A, TbSTT3B* and*TbSTT3C* was used for blotting. *TbSTT3A* is lacking after gene replacement, in agreement with the PCR analysis. However, the ectopic *TbSTT3A* is not observable and a new band at 2.9 kb was seen in the mutant cell line.

Therefore, new gene replacement trials were performed and genomic DNA was extracted from clonal cultures and once again screened by PCR. The same forward primer was used, however, the reverse primer anneals inside the BSD gene. In this way a PCR product of 0.85 kb will only be produced if BSD has replaced *TbSTT3A* in the

STT3 locus. Figure 3.1.14 displays a band at 0.85 kb, indicating that *TbSTT3A* has again been replaced.

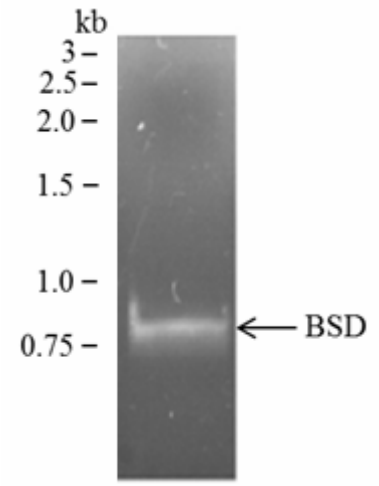


Figure 3.1.14.PCR screen of a *Tb*STT3A KO clone, showing replacement of *Tb*STT3A by BSD. PCR was performed using a forward primer that anneals upstream of the gene replacement construct and a reverse primer that anneals inside the BSD gene. The band at 0.85 kb indicates that BSD has replaced *TbSTT3A* in the *STT3* locus.

Positive cells were followed up by Southern blot analysis. This time a new digestion pattern using a different panel of restriction enzymes (SalI, PstI and BglII) was designed for the use with promiscuous *TbSTT3* probe (Figure 3.1.15).

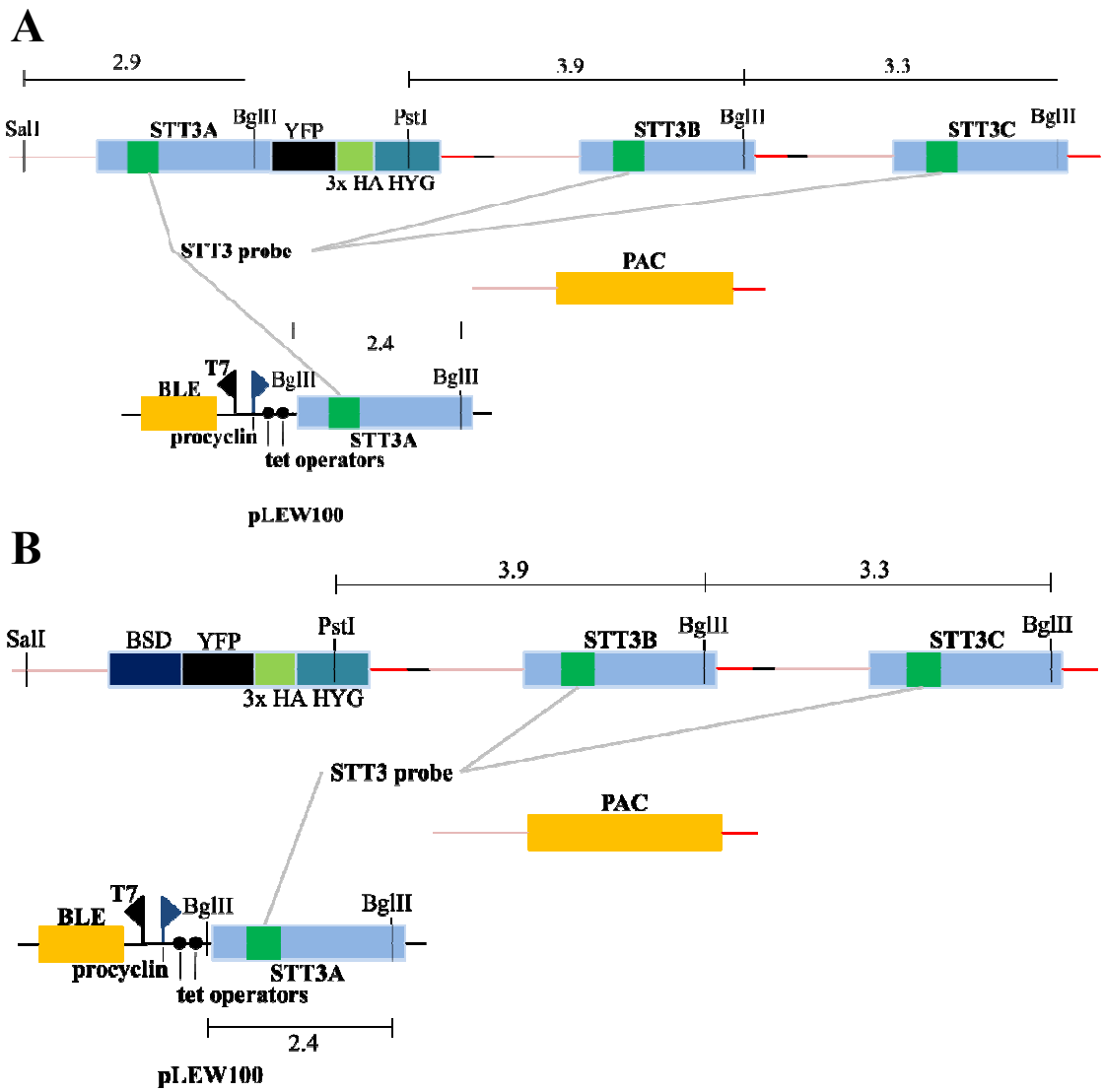


Figure 3.1.15. Digestion pattern for Southern blot visualising *TbSTT3A* gene replacement. The probe binds an identical sequence found in*TbSTT3A, TbSTT3B* and*TbSTT3C.***A:**By digesting with SalI, PstI and BglII four fragments of different sizes should be visualised if *TbSTT3A* is still present. **B:** The 2.9 kb band should be absent if BSD has replaced *TbSTT3A*.

With these restriction enzymes all the endogenous genes plus the ectopic *TbSTT3A* were visible at the correctly sized fragments with the parental cell line (Figure 3.1.16, lane 1) and, once again, loss of *Tb*STT3A was evident in the knock out clones. (Figure 3.1.16, lane 2). Nevertheless, similar to the previous Southern blot, new weak bands were

visible after gene replacement. This might be an indication of genomic rearrangement promoted by survival pressure during gene replacement, which has been published previously (Castillo-Acosta et al., 2013). It might be possible that genomic rearrangement will promote chimeric genes to retain the *Tb*STT3A function during gene replacement attempts. This could explain the weaker bands seen in lane 2. If the probing sequence is rearranged, different sized bands might be seen. Furthermore, less probe might bind to a rearranged sequence, yielding weaker bands.

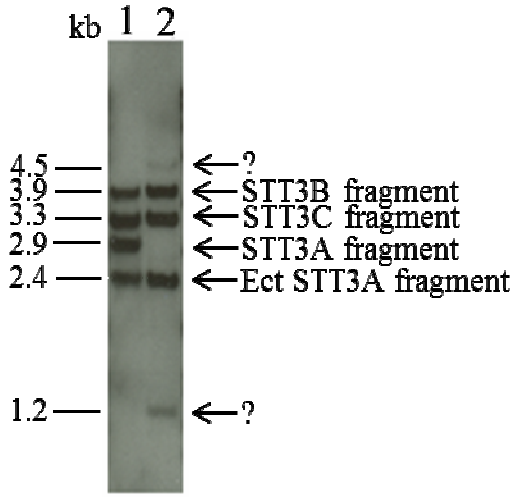


Figure 3.1.16. Southern blot visualising the loss of *TbSTT3A*.**Lane 1**: Genomic DNA (5 µg) from parental cell line after SalI, PstI and BglII digestion. **Lane 2**:Genomic DNA (5 µg) from transfected cells after SalI, PstI and BglII digestion. A promiscuous *STT3*probe binding an identical sequence found in*TbSTT3A, TbSTT3B* and*TbSTT3C* was used for blotting. *TbSTT3A* is missing after gene replacement, in agreement with the PCR analysis and previous Southern blot analysis. The ectopic *TbSTT3A* is seen at 2.4 kb. However, two new weak bands were seen after *TbSTT3A* gene replacement, at 1.2 and 4.5 kb.

As seen in Figure 3.1.17, the mutant cells showed no growth phenotype. To analyse if *Tb*STT3A function was still present, soluble VSGs from cells growing without tetracycline were again assessed by EndoH and PNGaseF digestion and analysed by SDS–PAGE and Coomassie blue staining (Figure 3.1.18).

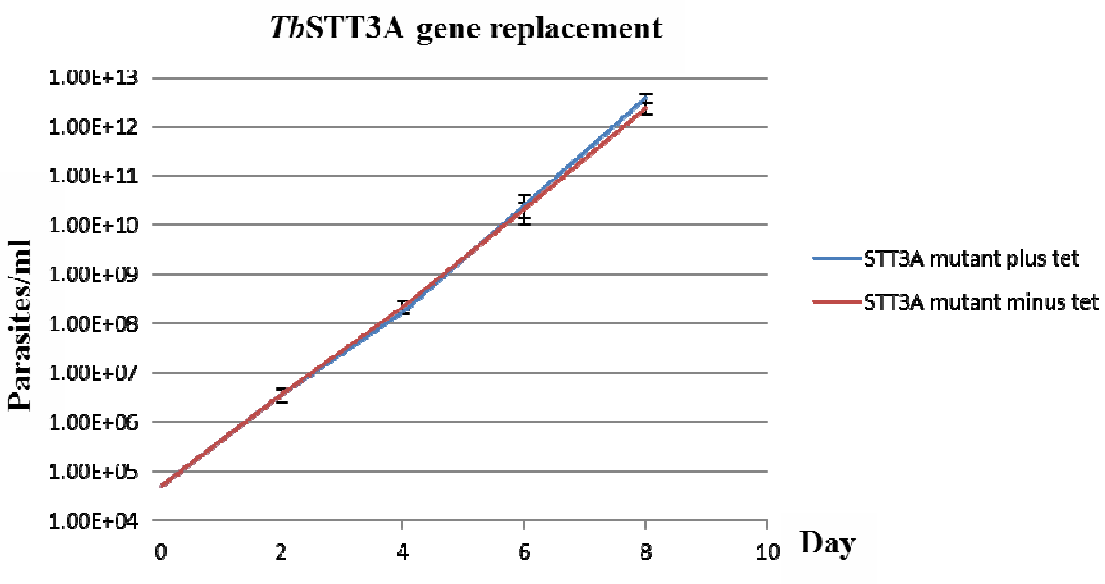


Figure 3.1.17.*Tb*STT3A mutant cells show no growth phenotype in culture. Cell numbers were counted in triplicate for eight days and cells were split when required. Cells grown in permissive and non-permissive conditions are seen in blue and red, respectively.

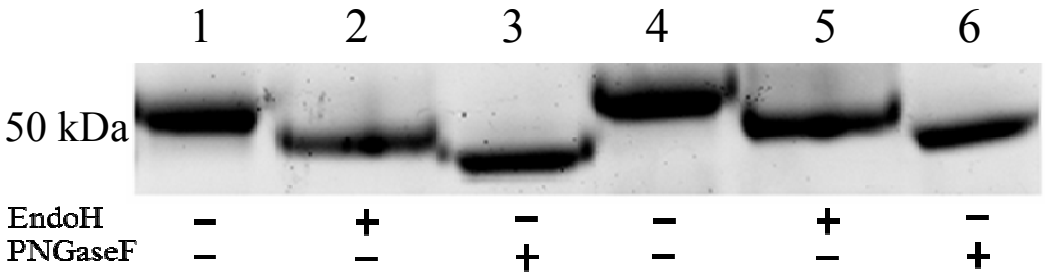


Figure 3.1.18.SDS-PAGE Coomassie stained gel visualising sVSG from the *Tb*STT3A gene replacement cell line and a wild type control. EndoH cleaves triantennary glycans whereas PNGaseF digests both tri- and biantennary glycans. **Lane 1-3**: VSG221 from wild type cell line after no (1), EndoH (2) or PNGaseF (3) treatment. **Lane 4-6**: VSG221 from *in situ* tagged cell line after no (1), EndoH (2) or PNGaseF (3) treatment. The digestion pattern indicates that $Man_5GlcNAc_2$ from the $Man_5GlcNAc_2$-PP-Dol precursor has been transferred to the asparagine at position 263 and, hence, that the function of*Tb*STT3A is still present.

The weight reduction seen after PNGaseF digestion indicates that a biantennary glycan has been cleaved suggesting $Man_5GlcNAc_2$is still being transferred from the $Man_5GlcNAc_2$-PP-Dol precursor to the asparagine at position 263. Therefore, altogether, the data suggest that the mutant cell line is not a *Tb*STT3A conditional null. Instead, it is likely that the function of *TbSTT3A* was preserved by genomic rearrangement in order to sustain cell growth during gene replacement, even in the presence of an expressed ectopic copy of the gene. This suggests that *Tb*STT3A is essential *in vitro*, though clearly further work is required to understand the putative genomic rearrangements that occur on trying to replace the second *Tb*STT3A allele.

3.1.4 Localisation studies of *Tb*STT3A

OST has previously been shown to interact with other ER membrane proteins (Wilson et al., 2005)and therefore the*T. brucei* enzyme has been presumed to also be integrated in the ER membrane. However, to date, no experimental data has demonstrated this. In order to investigate if *Tb*STT3A could be localised in *T. brucei*, experiments were performed with the *Tb*STT3A-YFP-HA_3 tagged cell line.

The genetically modified cells were prepared according to the protocol provided in Materials and Methods. Briefly, the cells were fixed in 4 % paraformaldehyde and plated on coverslips followed by washing, permeabilization and blocking. Subsequently, the coverslips were incubated with anti-Bip (ER marker) together with either anti-GFP or anti-HA (Table 3.1.1) and washed before microscopy analysis was performed. The results are seen in Figure 3.1.19.

Table 3.1.1. List of antibody combinations used for the *Tb*STT3A-YFP-HA_3 localisation studies.

Antibody combination	Dilution
α-GFP/α-Bip	1:400/1:2000
α-GFP/α-Bip	1:400/1:1000
α-GFP/α-Bip	1:300/1:2000
α-GFP/α-Bip	1:300/1:1000
α-HA/α-Bip	1:290/1:2000
α-HA/α-Bip	1:290/1:1000
α-HA/α-Bip	1:580/1:2000
α-HA/α-Bip	1:580/1:1000

Of the antibody combinations seen in Table 3.1.1, the one that showed the best results was anti-Bip and anti-HA with 1:2000 and 1:580 dilutions. It was not surprising that the anti-GFP gave no positive results since the tagged *Tb*STT3A was not detected using this antibody for Western blotting. The signal for both *Tb*Bip and the *in situ* tagged *Tb*STT3A was weak and further improvements will have to be carried out, nonetheless, *Tb*STT3A-YFP-HA_3 seems to co-localise with *Tb*Bip in Figure 3.19 (A, B and C). In A, one can see *Tb*STT3A highlighting the perinuclear ring, indicative of ER localisation.

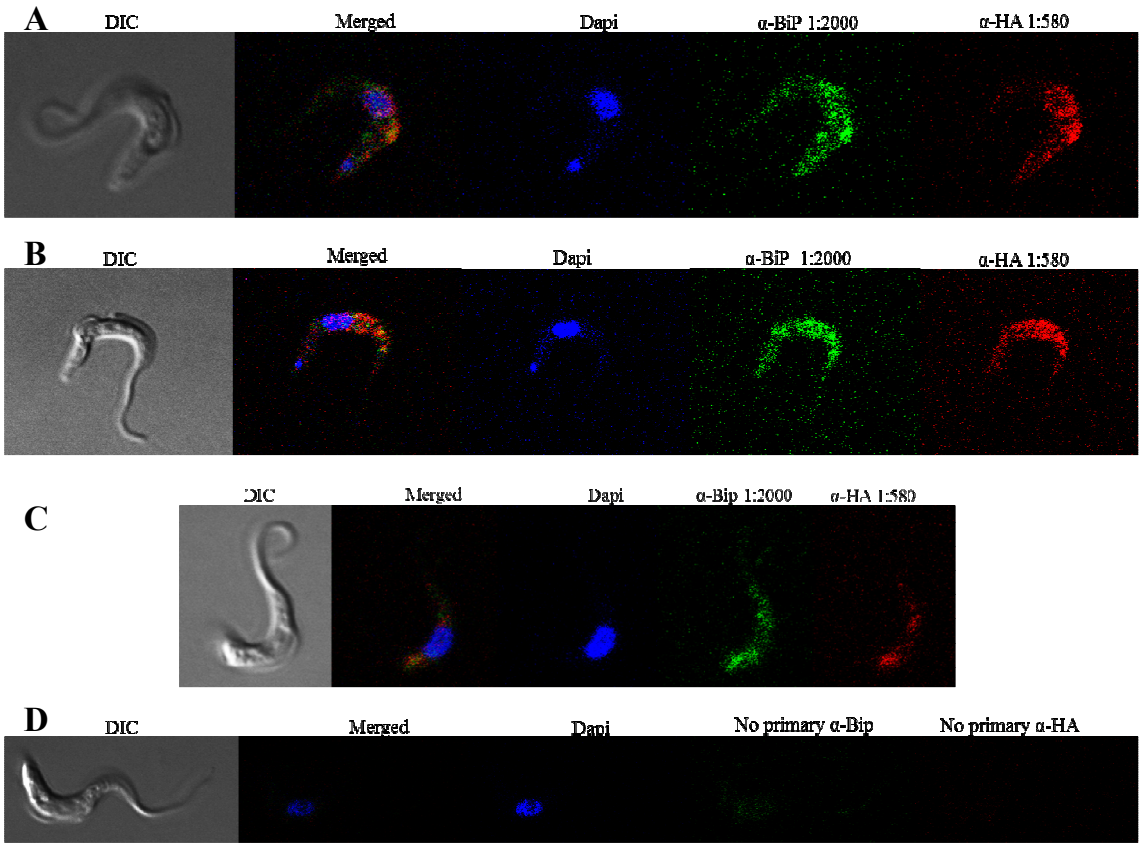


Figure 3.1.19. Localisation studies of *Tb*STT3A-YFP-HA$_3$. Cells were stained with Dapi (blue channel), α-Bip and Alexa 488-conjugated anti-mouse antibody (green channel) and α-HA and Alexa 594-conjugated anti-rabbit antibody (red channel). Although the images were not optimal, co-localisation of ER membrane protein *Tb*Bip and the YFP-HA$_3$ tagged T*b*STT3A was observed in **A**, **B** and **C**. In **D**, a negative control using no primary antibodies (α-Bip and α-HA) gives no signal.

The weak signal for *Tb*Bip can possibly be explained by an old aliquot of the antibody. The similarly weak signal for *Tb*STT3A-YFP-HA$_3$ might be due to low abundance of the proteinin combination with the large size of the ER. However, the results seen in Figure 3.19 suggested that *Tb*STT3A-YFP-HA$_3$ co-localises with *Tb*Bip. More optimisation will be needed to obtain clearer data proving that the *in situ* tagged *Tb*STT3A is localised in the ER membrane in *T. brucei*.

3.2 The search for novel *Tb*STT3A binding partners

Compared to the multiprotein complexes in higher eukaryotes, *T. brucei* oligosaccharyltransferase (OST) is suggested to comprise a single protein, the catalytic subunit STT3. This hypothesis originatedbecause no homologues were found in bioinformatic searches of the *T. brucei*genome, indicating that it lacks all OST subunit homologues except for the catalytic STT3 subunit, for which there are three complete paralogous genes. (Berriman, 2005; Izquierdo et al., 2009a). However, it has been shown before, for example for the GPI biosynthetic machinery that, in some cases structural (but not sequence) homology can exist (Nagamune et al., 2003). To investigate if the oligosaccharyltransferase in *T. brucei* has any such structural homologue subunits and/or novel binding partners, immunoprecipitation and blue native gel electrophoresis experiments were performed using *in situ* tagged *Tb*STT3A.

3.2.1 *Tb*STT3A-YFP-HA_3 immunoprecipitation

The*in situ* tag introduced in the previous chapter was meant to enable immunolocalisation studies and facilitate pull-down experiments to investigate potential *Tb*STT3A binding partners. The C-terminal tag consists of a eukaryotic YFP followed by a three hemagglutinin (HA_3) epitopes. In between the two tags, a TEV site was introduced using the primers found in Table 2.1. To detect the tagged *Tb*STT3A by Western blotting, $1*10^9$ cells were lysed before immunoprecipitation with either anti-GFP or anti-HA according to the protocol in section 2.4.5 of Materials and Methods. Membrane proteins are known to be challenging to pull down and heating to 100 °C has been suggested to promote aggregation. To test this, cells were lysed in 1 % SDS in the

presence of 0.1 M DTT and protease inhibitors, followed by heating up to 100 °Cor 50 °C for 10 minutes. The lysates were subsequently diluted to 0.03 % SDS to facilitate immunoprecipitation. Figure 3.2.1 shows that *Tb*STT3A was detected after anti-HA pull down following heating to both 50 °Cand 100 °C (lane 1 and 2). However, *Tb*STT3A was not detected after immunoprecipitation by anti-GFP antibodies (lane 4 and 5). A wild-type cell line was used as a negative control (lane 3 and 6).

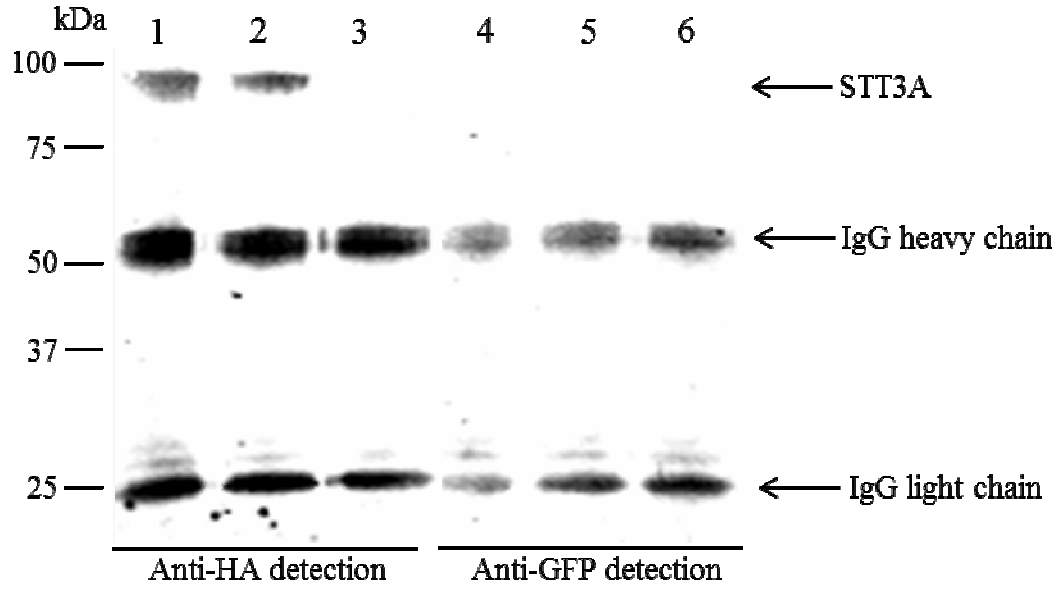


Figure 3.2.1. Western blot results after anti-HA and anti-GFP immunoprecipitation. Lane 1 and 2: Anti-HA immunoprecipitation and detection after heating to 50°C (lane 1) or 100 °C (lane 2) during cell lysis. *Tb*STT3A-YFP-HA_3 is detected at approximately 95 kDa. Lane 4 and 5: Anti-GFP immunoprecipitation and detection after heating to 50°C (lane 4) or 100 °C (lane 5) during cell lysis. The tagged *Tb*STT3A was not detectable when using anti-GFP. In lane 3 and 6, a wild-type cell line was used as a negative control, lysed at 100 °C and detected by anti-HA (lane 3) or anti-GFP (lane 6).

Using anti-HA antibody for both immunoprecipitation and detection, the tagged *Tb*STT3A-YFP-HA_3 fusion protein was detected at approximately 95 kDa. The theoretical molecular weight for YFP and HA_3 tagged *Tb*STT3A is 117 kDa and the smaller molecular weight detected is in agreement with previously published results (Izquierdo et al., 2009a) where untagged *Tb*STT3A was detected at a molecular weight at about 70 kDa (theoretical molecular weight 89 kDa). Since the tagged *Tb*STT3A was not detectable by anti-GFP immunoprecipitation, concerns were raised about the YFP tag. In YFP, a threonine has been mutated into a tryptophan compared to the GFP tag, and GFP antibodies should still be able to bind YFP. Furthermore, since the HA_3 tag is found downstream of the YFP, no nonsense or stop codon mutation can have occurred. Additionally, since the proteins are denatured prior to Western blotting, YFP misfolding should not be a major issue. Therefore, we assumed that anti-GFP immunoprecipitation of *Tb*STT3A was below the detection limit when starting with $1*10^9$ cells. Since the HA_3 epitope functioned well when used to pull down *Tb*STT3A we decided to remove the YFP tag. This should also improve the chances of detecting potential binding partners since the 27 kDa YFP tag might have blocked potential interactions.

3.2.2 *In situ* tagging *Tb*STT3A with HA_3

Similarly to the previous *in situ* tagging strategy of *Tb*STT3A, the C-terminus of the gene (excluding the stop codon) and the 3' UTR directly downstream of *Tb*STT3A were cloned into the pMOTag4H plasmid (Figure 3.2.2). This plasmid only introduces a HA_3 tag, which was directed to the C-terminus, under HYG selection.

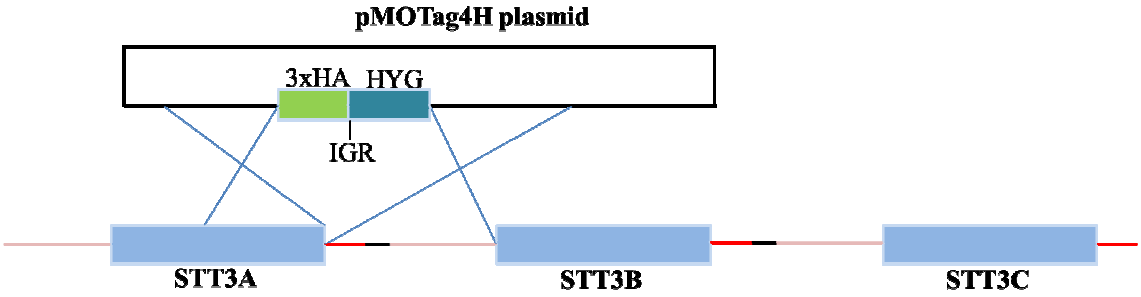


Figure 3.2.2. Strategy for the generation of the pMOTag4H *in situ* tagging construct. The C-terminus of *TbSTT3A* and the 3' UTR was PCR amplified and cloned into the pMOTaG4H plasmid.

The plasmid was linearized before being transfected into the *Tb*STT3A,B,C +/- cell line. No overexpressing ectopic copy was needed during electroporation since previous data suggested that these tags do not interfere with the function of the protein. The HA_3 epitope will fuse to the C-terminus of *Tb*STT3A through homologous recombination (Figure 3.2.3) and cells were subsequently selected by adding HYG to the media.

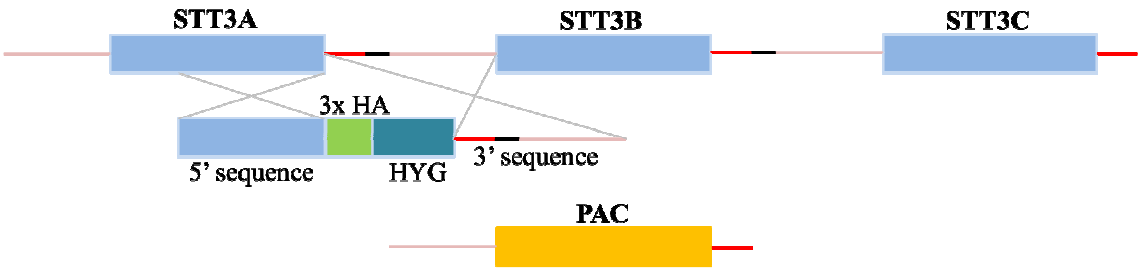


Figure 3.2.3. Strategy for the HA_3*in situ* tagging of *TbSTT3A*.The C-terminus of *TbSTT3A* will befused to an HA_3 tag under HYG selection. The tag was introduced to the *TbSTT3A,B,C* +/- heterozygote cell line.

Following transfection, recovered cells were made clonal and the genotype was analysed by Southern blotting. When probing against HYG following PstI digestion

(Figure 3.2.4), specific bands of the correct size were seen, confirming that *Tb*STT3A has been correctly *in situ* tagged with an HA_3 epitope (Figure 3.2.5).

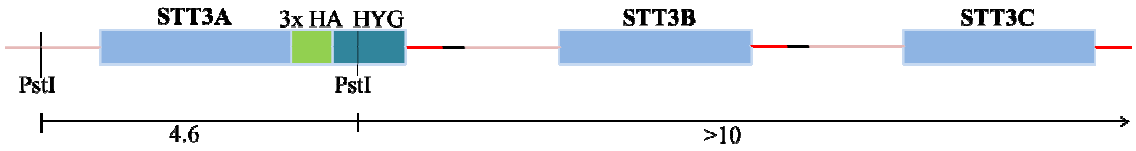


Figure 3.2.4. Predicted digestion pattern for Southern blot analysis of *in situ* tagged *TbSTT3A*. When probing against HYG following PstI digestion, two bands should be observable at 4.6 and >10 kb.

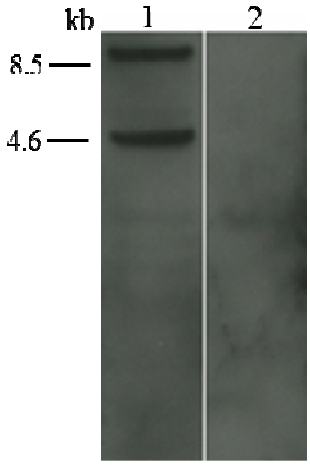


Figure 3.2.5. Southern blot visualising HA_3*in situ* tagging of *TbSTT3A*.Lane 1: Genomic DNA (5 µg) from transfected cells after PstI digestion. Lane 2: Genomic DNA (5 µg) from parental cell line after PstI digestion. The PstI digestion resulted in two predicted bands, 4.6 and >10Kb, indicating that only *Tb*STT3A has been *in situ* tagged. A HYG probe was used for the analysis.

The *Tb*STT3A-HA_3 cell line showed no growth phenotype (data not shown). To verify that the enzyme function was not impaired by the tag, sVSG was purified and N-glycans

were digested with EndoH and PNGaseF, as described in Material and Methods (and briefly in section 3.1.2). In Figure 3.2.6, the same reduction in molecular weight was seen after EndoH and PNGaseF digestion as compared to a wild-type cell line, confirming that the *Tb*STT3A-HA_3 is still able to transfer $Man_5GlcNAc_2$and that the functions have remained intact during *in situ* tagging.

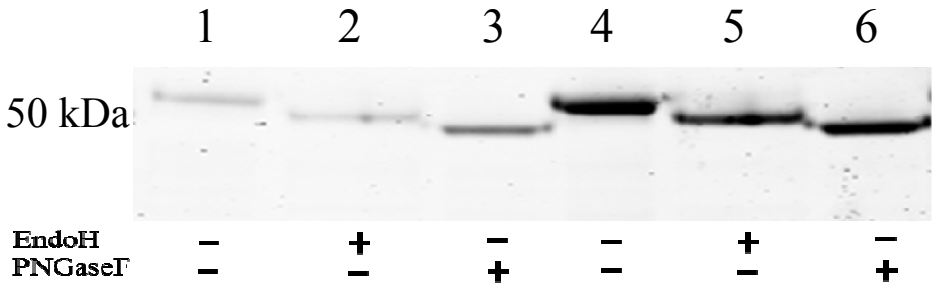


Figure 3.2.6.SDS-PAGE Coomassie stained gel visualising sVSG from the pMOTag4H transfected *Tb*STT3A-HA_3*in situ* tagged cell line and a wild-type control. EndoH cleaves off triantennary glycans whereas PNGaseF digests both tri- and biantennary glycans. The digestion pattern indicates that $Man_5GlcNAc_2$ from the $Man_5GlcNAc_2$-PP-Dol precursor has been transferred to the asparagine at position 263 and, hence, that the function of the tagged *Tb*STT3A was unaffected.

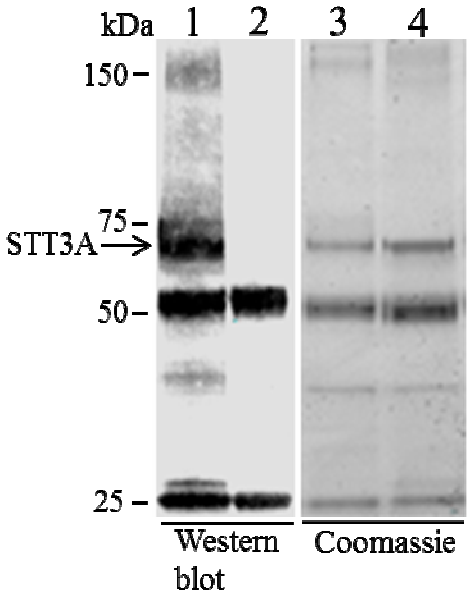
3.2.3 *Tb*STT3A-HA_3 immunoprecipitation under stringent conditions

After verifying that *Tb*STT3A had been correctly tagged with an HA_3 epitope and that function was conserved, new attempts to immunoprecipitate the protein were initiated. With the new *Tb*STT3A-HA_3 cell line, the protein was detected after direct Western blotting, loading the cell equivalent of $5*10^6$ cells on the gel (data not shown). Visualised in Figure 3.2.7 are the Western blot and Coomassie stained gel after pulling down *Tb*STT3A with an anti-HA antibody under stringent conditions, as described in

section 2.4.5. Briefly, $5*10^8$ cells were lysed at 100 °C in 1 % SDS and 0.1 M DTT in the presence of protease inhibitors and the lysate was later diluted to 0.03 % SDS before being incubated with anti-HA mouse antibody. Finally, magnetic beads binding the antibodies were added and subsequently captured with a magnet and washed extensively. Bound proteins were then eluted before being loaded on a gel. After Western blotting with anti-HA, *Tb*STT3A was detected at just below 75 kDa in the *Tb*STT3A-HA_3 cell lysate (lane 1). No band was seen at the same molecular weight when using a wild-type cell lysate as a negative control. Lanes 3 and 4 show the same two cell lysates (*Tb*STT3A-HA3 and wild-type) separated on a gel and Coomassie stained. Loading appeared equal and the bands seen at 25 and 50 kDa are the light and heavy chain of the mouse anti-HA used in the pull downs. The secondary anti-mouse antibody detects both these IgG fragments and potential anti-HA mouse antibodies bound to HA epitopes.

Figure 3.2.7.(Next page) Western blotted and Coomassie stained gels after separation of cell lysate from *Tb*STT3A-HA_3 and wild-type cell lines. Lanes 1 and 2 show *Tb*STT3A-HA_3 and wild-type lysate following Western blotting using an anti-HA antibody. *Tb*STT3A is detected below 75 kDa in lane 1, however, a band at the same molecular weight was not seen in the wild-type sample shown in lane 2. Lanes 3 and 4 show the same cell line lysates with Coomassie staining. No band was seen at the same molecular weight in the control shown in lane 4. Bands seen at 25 and 50 kDa in all lanes are the light and the heavy chain of the antibody used in the immunoprecipitation.

Figure 3.2.7



3.2.4 Optimising mild detergent conditions for immunoprecipitation

Following the pull down of *Tb*STT3A in stringent conditions, the next step was immunoprecipitation of the protein in a native state, facilitating the search for potential binding partners. Digitonin is a detergent that effectively solubilises lipids and hence a common choice for immunoprecipitation of membrane proteins. Most published data on solubilising membrane proteins for pull down uses 1 % digitonin during cell lysis (Oeljeklaus et al., 2012; Roboti and High, 2012b; Singha et al., 2012). To analyse how the concentration of digitonin affects interacting proteins, a test experiment was set up using VSG221 as a reporter protein. Native VSG is a homodimer with a molecular weight of approximately 100 kDa (Cross, 1975b). The VSGs were diluted in 1.0, 0.8, 0.6, 0.4, 0.2 and 0.0 % digitonin before being loaded on a blue native gel (Materials and Methods, 2.4.7) As seen in Figure 3.2.8, the VSG dimer (100 kDa) is the dominant form when using 0, 0.2 and 0.4 % digitonin. By diluting the dimer in 0.6% digitonin, half of the dimers disassociated into their monomeric state (50 kDa). When using even higher concentrations of digitonin the monomeric state is the predominant one seen.

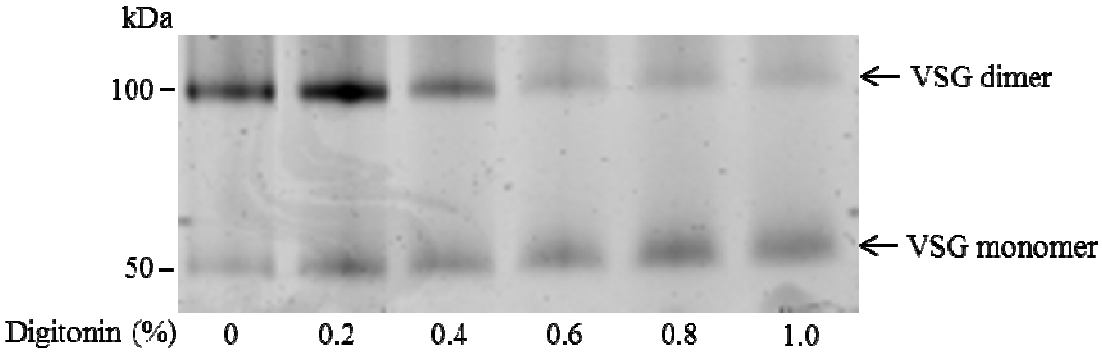


Figure 3.2.8. Blue native gel showing the quaternary structure of VSG after being diluted in different concentrations of digitonin. When no digitonin detergent is added or when diluting to 0.2 or 0.4 %, the dimer is the major structural state. When diluting to 0.8 or 1.0 % digitonin, the monomeric state was predominant. The same amount of dimer and monomer is detected after diluting VSG in 0.6 % digitonin.

3.2.5 *Tb*STT3A-HA$_3$ immunoprecipitation under mild conditions

After showing that use of more than 0.8% digitonin significantly disrupted the VSG dimer, we decided to test lysing *T. brucei* cells in 0.5 % digitonin (described in section 2.4.5). Briefly, $5*10^8$ cells were lysed in 0.5 % digitonin on ice for 30 minutes, followed by incubating the cell lysate with anti-HA mouse antibody. Magnetic beads, which will bind to the antibodies, were subsequently added and following capture of the magnetic beads, samples were washed and the bound proteins were eluted. The samples were subsequently loaded on a gel and Western blotted using a rabbit anti-HA antibody. The results are seen in Figure 3.2.9. *Tb*STT3A was detected at just below 75 kDa, in agreement with the pull down in stringent conditions. No bands were visible in the wild-type sample. Below the Western blot, a Ponceau S stained loading control verifies that the samples were equally loaded.

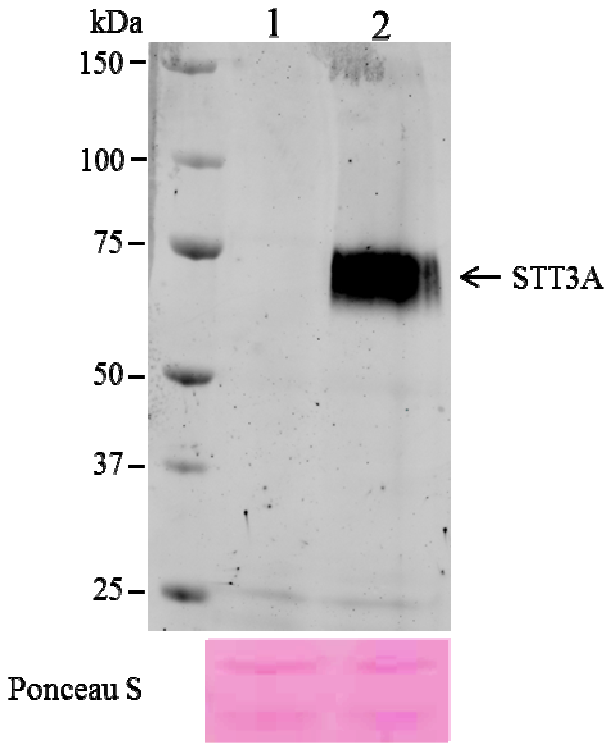


Figure 3.2.9. Western blot result after immunoprecipitation of *Tb*STT3A-HA_3 from digitonin lysate.**Lane 1**: Cell lysate from wild-type cells. **Lane 2**: Cell lysate from *Tb*STT3A-HA_3 cells. *Tb*STT3A was detected at ~70 kDa in lane 2. A Ponceau S stained loading control is seen under the Western blot.

3.2.6 *Tb*STT3A-HA_3 blue native gel electrophoresis

Previous results had shown that *Tb*STT3A could be detected at ~70 kDa by Western blotting following immunoprecipitation with anti-HA antibody in both stringent and mild lysing conditions. As a first step to experimentally investigatingwhether*Tb*STT3A has any novel binding partners, blue native gel electrophoresis analysis was performed with the *Tb*STT3A-HA_3 cell line. Cells were lysed in 0.5 % digitonin for 30 minutes on ice before being loaded on a NativePAGE (Invitrogen) gel and finally blotted on a PVDF membrane. Anti-HA antibody was used for detection. In Figure 3.2.10, it is clear that no band is visible at approximately 75 kDa as had been seen under reducing

conditions. However, a big smear, seen only in the *Tb*STT3A-HA_3 cell line in lane 2 was detected between 700 and 1200 kDa, suggesting that *Tb*STT3A might be in a large complex.

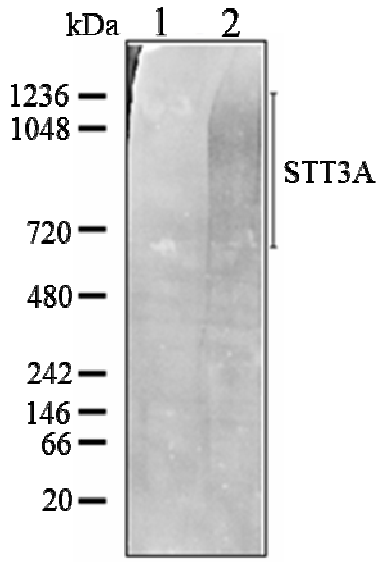


Figure 3.2.10. Western blot results after running blue native gel with *Tb*STT3A-HA_3 and wild-type cell lysates. Lane 1: No band is seen when using wild-type cell lysate. Lane 2: A big smear is detected between 700 and 1200 kDa when using *Tb*STT3A-HA_3 cell lysate.

3.2.7 *Tb*STT3A-HA_3 SILAC immunoprecipitation

Since the results from the blue native gel electrophoresis suggested *Tb*STT3A to be in a large complex, we decided to carry out immunoprecipitation experiments using stable isotope labelling in cell culture (SILAC, section 2.5). By doing this, we should be able to distinguish genuine *Tb*STT3A binding partners from contaminants according to peptide isotopic signatures. For the experiment, wild-type and transgenic parasites expressing *Tb*STT3A-HA_3 were grown in parallel and under identical conditions for eight cell divisions. However, the transgenic *Tb*STT3A-HA_3 cell line was grown in

"heavy medium" containing stable isotope-labelled Lys and Arg (R_6K_4), whereas the wild-type cells were grown in "light medium" containing unlabelled Lys and Arg (R_0K_0). The transgenic *Tb*STT3A-HA$_3$ parasites and the wild-type cells were harvested, washed, counted, mixed together in a 1:1 ratio, and lysed in 0.5 % digitonin. These lysates, containing equal amounts of steady-state labelled and unlabelled proteins, were used to capture the heavy-labelled *Tb*STT3A-HA$_3$ proteins on anti-HA magnetic beads for proteomic analysis. In this manner, specifically enriched *Tb*STT3A-HA$_3$ proteins and potential associated proteins can be distinguished from nonspecific contaminant proteins by the isotope ratios of their tryptic peptides. Thus, *Tb*STT3A-HA$_3$ and true associated protein peptides will have high heavy/light isotope ratios, whereas contaminant proteins will have heavy/light isotope ratios close to 1 (Figure 3.2.11).

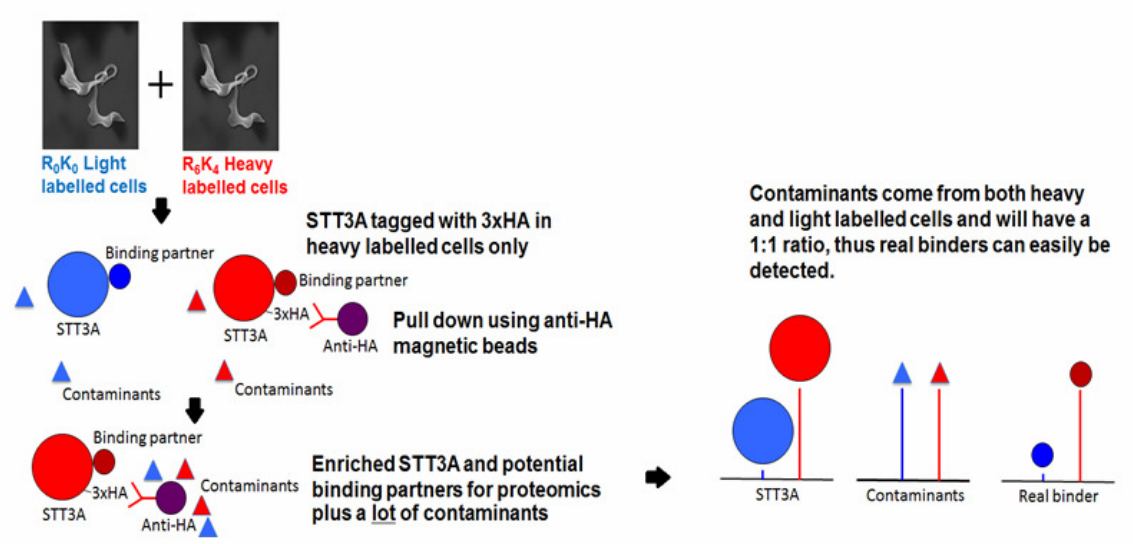


Figure 3.2.11. Overview of the SILAC immunoprecipitation experiment. Wild-type bloodstream form cells were grown in light (R_0K_0) medium, and cells expressing *Tb*STT3A-HA$_3$ were labelled with heavy (R_6K_4) Lys and Arg. The cells were mixed 1:1 and *Tb*STT3A-HA$_3$ was enriched by affinity-selection on anti-HA magnetic beads. Peptides from *Tb*STT3A-HA$_3$ and genuine binding proteins have high heavy/light isotope ratios, whereas those from contaminants have ratios close to 1:1.

The data set from the experiment was used to search a *T. brucei* predicted protein database using MaxQuant software. Each protein identification was displayed on a plot (Figure 3.2.12) of the Log10 value of the intensities of the unique peptides of that protein (y axis) and the Log2 value of the heavy to light isotope ratios of the same peptides (x axis).

*Tb*STT3A (bait) had the highest heavy/light ratio (14.6) closely followed by *Tb*STT3B (9.6) (Table 3.2.1). After these two proteins, the ratio drops to around 1.5. Even though a few of these proteins had been significantly enriched (Figure 3.2.12), none of them were localised in the ER and were thus categorised as possible contaminants. Conclusively, *Tb*STT3A and *Tb*STT3B were highly enriched (Figure 3.2.12) compared to the rest of the proteins detected in the experiment, suggesting that *Tb*STT3A and *Tb*STT3B form an oligosaccharyltransferase complex with no other subunits in *T. brucei*.

Table 3.2.1. Top five hits in the *Tb*STT3A-HA_3 SILAC IP.

Protein ID	Description	Intensity	Heavy/light norm.
Tb927.5.890	*Tb*STT3A oligosaccharyltransferase	38682000	14.6
Tb927.5.900	*Tb*STT3B oligosaccharyltransferase	8866190	9.6
Tb927.10.4560	Elongation factor 2	469935	1.4
Tb927.10.14140	Pyruvate kinase 1 (PYK1)	369950	1.3
Tb927.8.3530	Glycerol-3-phosphate dehydrogenase [NAD]	1918180	1.3

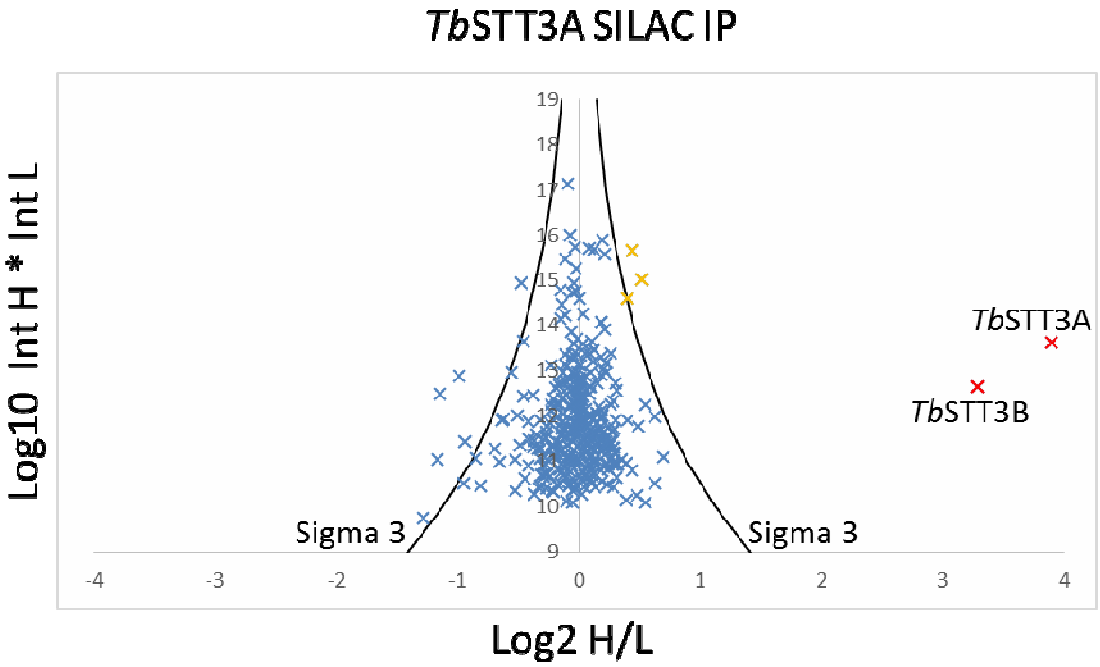


Figure 3.2.12. Plot of the *Tb*STT3A-HA_3 immunoprecipitation proteomics data. Each protein identification is represented by a point plotted as the Log2 of their heavy-to-light isotope ratio (x axis) versus the Log10 value of the intensities of the peptides belonging to each protein (y axis). The black curves represent three standard deviations from the mean. Proteins plotted in orangehave a heavy-to-lightratio above the sigma 3 cut off and are significantly enriched. *Tb*STT3A (bait) and *Tb*STT3B (annutated) were shown to be highly enriched and are highlighted in red.

3.2.8 *In situ* tagging TbSTT3B with MYC_3

The result from the SILAC immunoprecipitation experiment suggested that *Tb*STT3A is in a complex with *Tb*STT3B and that no other novel subunits are part of the *T. brucei* oligosaccharyltransferase. To address this hypothesis, we wanted to *in situ* tag *Tb*STT3B with a MYC_3 epitope, followed by pull-out experiments of the tagged protein. By doing this we should be able to detect *Tb*STT3A when pulling down

*Tb*STT3B (and *vice versa*) by Western blotting. Furthermore, by performing pull-out experiment using anti-MYC both *Tb*STT3B (bait) and *Tb*STT3A should be enriched. *Tb*STT3B was *in situ* tagged as previously described. However, the in *Tb*STT3B *situ tagging* construct (called pMOTag4M*) was ordered from Genescript (Appendix). The construct incorporates a MYC_3 tag under blasticidin (BSD) selection and was designed to contain ~1000 bp of the C-teminus of the *Tb*STT3B and the 3' UTR downstream of the gene (Figure 3.2.13). The C-terminus of *Tb*STT3B is highly similar to the one of *Tb*STT3C, but since the 3' UTR is only found downstream of *Tb*STT3A and *Tb*STT3B, the tag should only fuse to the C-terminus of *Tb*STT3B.

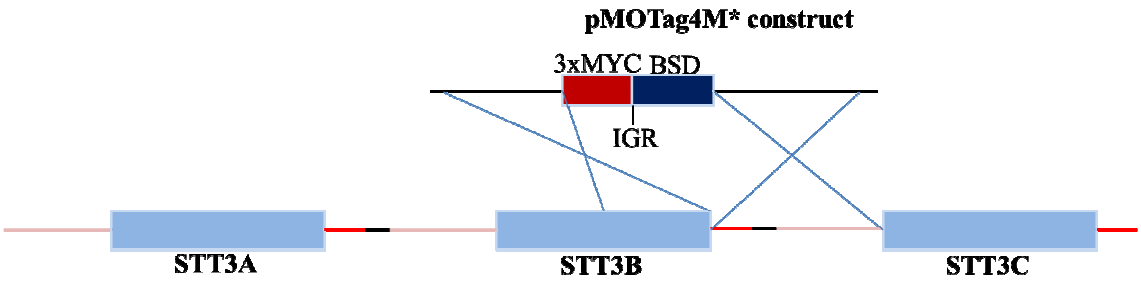


Figure 3.2.13. Schematic diagram of the pMOTag4M* construct that was ordered from Genescript to facilitate *in situ* tagging of *Tb*STT3B. The construct will specifically tag endogenous *Tb*STT3B with a MYC_3 epitope under BSD selection.

The plasmid was transfected into the *Tb*STT3A,B,C +/- ; *Tb*STT3A-HA_3 cell line. The MYC_3 epitope should fuse to the C-terminus of *Tb*STT3B through homologous recombination (Figure 3.2.14) and cells were subsequently selected by adding BSD to the media.

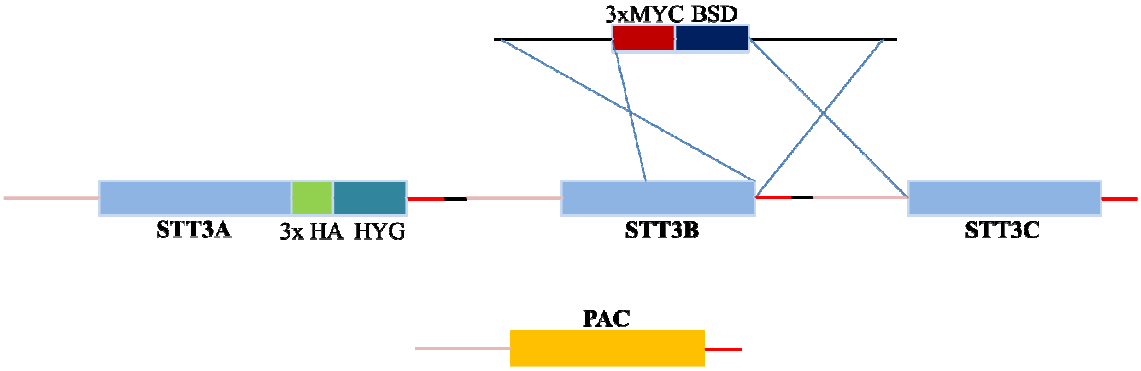


Figure 3.2.14. Strategy for the MYC_3*in situ* tagging of *TbSTT3B.*The C-terminus of *TbSTT3B* will befused to an MYC_3 tag under BSD selection. The tag was introduced to the *TbSTT3A,B,C* +/- ; *Tb*STT3A-HA_3 cell line.

Following transfection, recovered cells were made clonal and tagged *Tb*STT3B was detected at the correct molecular weight by Western blot after anti-MYC immunoprecipitation in stringent conditions (data not shown), confirming that the cell line illustrated in Figure 3.2.15 had been established.

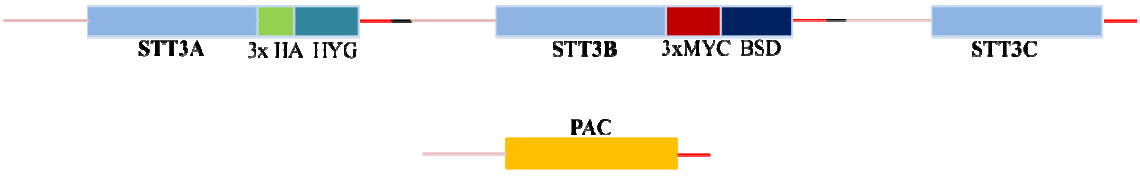


Figure 3.2.15. Doubly *in situ* tagged cell line established after electroporating the pMOTag4M* construct and selecting clonal cells with BSD.

3.2.9 *Tb*STT3B-MYC$_3$ and *Tb*STT3A-HA$_3$ co-immunoprecipitation under mild conditions

5*10^8 cells from the abovementioned cell line were harvested, washed and lysed in 0.5 % digitonin. *Tb*STT3A-HA$_3$ or *Tb*STT3B-MYC$_3$ were captured from the lysate using anti-HA or anti-MYC magnetic beads and subsequently the tagged proteins were detected by Western blotting using anti-HA or anti-MYC antibodies. Figure 3.2.16 visualises the results from the immunoprecipitations. Lane 1 to 4 are the results following anti-HA (lane 1 and 2) and anti-MYC (lane 3 and 4) pull down and anti-HA detection. The band seen just below the 75 kDa marker in lane 2 is *Tb*STT3A-HA$_3$ after anti-HA immunoprecipitation, verifying that the HA tagged protein has been pulled out from the lysate of the transgenic *Tb*STT3A-HA$_3$/*Tb*STT3B-MYC$_3$ cell line but not from wild-type cell lysate (lane 1). A weak band of the same size is seen in lane 4. This suggests that *Tb*STT3A-HA$_3$ has been co-immunoprecipitated with the MYC tagged *Tb*STT3B and been detected by the anti-HA used for Western blotting. No band at the same molecular weight was detected in the wild-type control in lane 3. Lane 5 to 8 are the results following anti-HA (lane 5 and 6) and anti-MYC (lane 7 and 8) pull down and anti-MYC detection. The band seen just belowthe 75 kDa marker in lane 8 is *Tb*STT3B-MYC$_3$ after anti-MYC immunoprecipitation, verifying that the MYC tagged protein has been pulled out from the lysate of the transgenic *Tb*STT3A-HA$_3$/*Tb*STT3B-MYC$_3$ cell line but not from wild-type cell lysate (lane 7). A band of the same molecular weight is seen in lane 6. This suggests that *Tb*STT3A-MYC$_3$ has been co-immunoprecipitated with the HA tagged *Tb*STT3A and been detected by the anti-MYC used for Western blotting. No band at the same molecular weight was detected in the wild-type control in lane 5. These results are in agreement with the SILAC *Tb*STT3A pull down, suggesting that *Tb*STT3A and *Tb*STT3B form an oligosaccharyltransferase complex in *T. brucei*.

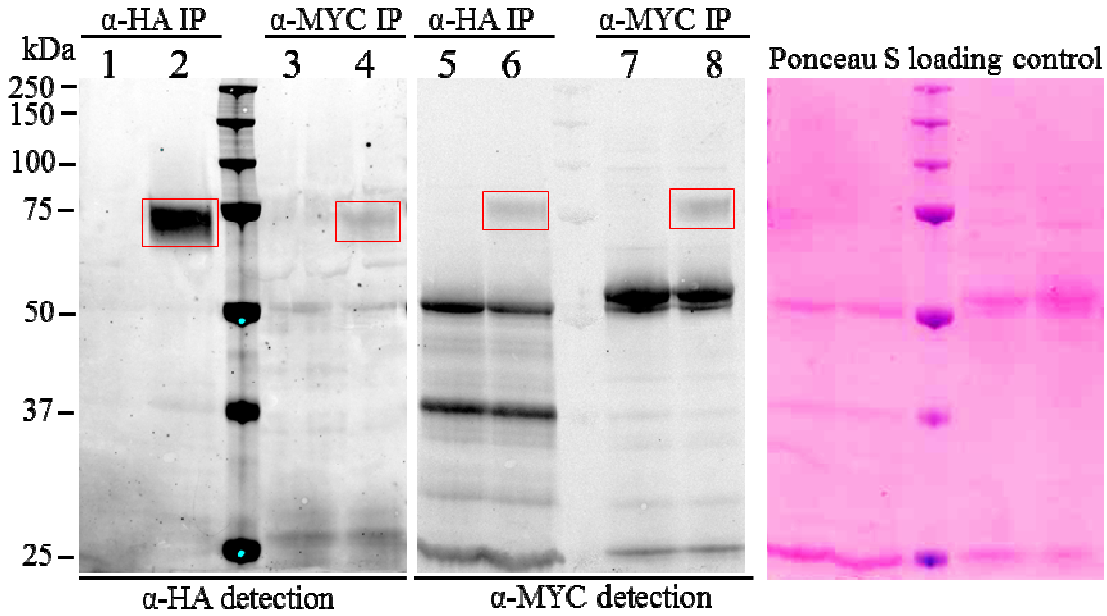


Figure 3.2.16. Western blot results after anti-HA and anti-MYC pull down. Anti-HA was used for detection of *Tb*STT3A-HA_3 in lane 1-4, whereas anti-MYC was used for detection of *TB*STT3B-MYC_3 in lane 5-8. Highlighted in red boxes are bands visualising *Tb*STT3A-HA_3 and *Tb*STT3B-MYC_3 (lane 2 and 4 and lane 6 and 8, respectively). No bands were detected after using wild-type cell lysate for immunoprecipitation (lane 1, 3, 5 and 7). The weak band seen highlighted in lane 4 following anti-MYC pull down confirms that some *Tb*STT3A-HA_3 co-immunoprecipitate with the MYC tagged *Tb*STT3B. The band highlighted in lane 6 following anti-HA pull down confirms that some *Tb*STT3B-MYC_3 co-immunoprecipitate with the HA tagged *Tb*STT3A. The Ponceau S stained membrane confirms equal loading.

Taken together, these data suggest that no novel subunits take part in *T. brucei* OST. However, *Tb*STT3A appears to be bound to *Tb*STT3B in the native state, thus forming a large multimeric OST complex together. To our knowledge, this is the first time these two catalytic subunits have been suggested to be in a complex with each other.

3.3 Acceptor peptide specificity of *Tb*STT3A and *Tb*STT3B

In 2005 Jones *et al* reported that *T. brucei* uses two precursor N-glycans, $Man_5GlcNAc_2$ and $Man_9GlcNAc_2$,unlike the single $Glc_3Man_9GlcNAc_2$ precursor used under normal conditions in most other eukaryotes(Jones et al., 2005). *Tb*STT3A is responsible for co-translational transfer of the biantennary $Man_5GlcNAc_2$ to specific N-glycosylation sequons, whereas *Tb*STT3B catalyses post-translational transfer of triantennary $Man_9GlcNAc_2$ to the remaining sterically accessible sequons(Izquierdo et al., 2009a). Through proteomic and bioinformatic studies, Izquierdo *et al.* also proposed that $Man_5GlcNAc_2$ is transferred to sequons in an acidic environment while $Man_9GlcNAc_2$ is transferred to remaining sequons in both neutral and basic amino acid surroundings(Izquierdo et al., 2009a). To improve our understanding of the acceptor peptide specificity in *T. brucei*, an *in vivo* assay was established. Using a glycosylation reporter protein, we could assess what type(s) of glycan(s) had been attached to a single N-linked glycosylation sequon on the protein, and thus elucidate which enzyme transferred that glycan, *Tb*STT3A or *Tb*STT3B (or both). Furthermore, the *in vivo* assay highlighted how the amino acids surrounding the sequon may affect oligosaccharyltransferase specificity in *T. brucei*.

3.3.1 Generation of *Tb*BipN N-linked glycosylation reporter protein

To establish an *in vivo* assay that could assess whether a biantennary $Man_5GlcNAc_2$ or a triantennary $Man_9GlcNAc_2$ glycan had been transferred by *Tb*STT3A or *Tb*STT3B to an N-linked glycosylation sequon, a reporter protein with a single N-X-S/T site was needed. By removing the ER retention sequon of ER membrane protein *Tb*Bip, the

smaller protein, *Tb*BipN, is transferred through the secretory pathway and secreted(Bangs et al., 1996). *Tb*BipN has no native N-linked glycosylation sequon, enabling introduction of a single known sequon, so this protein was chosen to be our glycosylation reporter. The *Tb*BipN protein coding DNA sequence was PCR amplified using previously published primers (Bangs et al., 1996) and subsequently cloned into a pLEW82 plasmid. The pLEW82 plasmid is a tetracycline-inducible high expression plasmid but with poor regulation. A single T7 promoter drives the expression of both the gene of interest and a phleomycin resistance gene. In addition, the plasmid has one tetracycline operator for regulation of the expression and two T7 terminators to prevent any T7 polymerase read-through of downstream genes.

The construct encodes the *Tb*BipN protein fused to the amino acid sequence PRAAAXXXXX**NAT**XXXXXQLAAA, followed by three HA tag sequences to facilitate Western blotting (Figure 3.3.1, A). The glycosylation sequon, **NAT,** is flanked by five variable amino acids on either side (X) and the PR and QL (proline/arginine and glutamine/leucine) amino acids are encoded by the AvrII and MfeI restriction sites that are used to introduce different experimental XXXXX**NAT**XXXXX sequences. The two AAA sequences (underlined) provide neutral linker sequences beyond the two experimentally variable sequences(XXXXX). This construct is incorporated into the rRNA spacer and following pheomycin selection, *Tb*BipN will be expressed with either a biantennary paucimannose glycan (transferred by *Tb*STT3A) or a triantennary oligomannose glycan (transferred by *Tb*STT3B) (Figure 3.3.1, B). It is also possible that both of these glycans can be transferred to the aspargine in the N-X-S/T sequon.

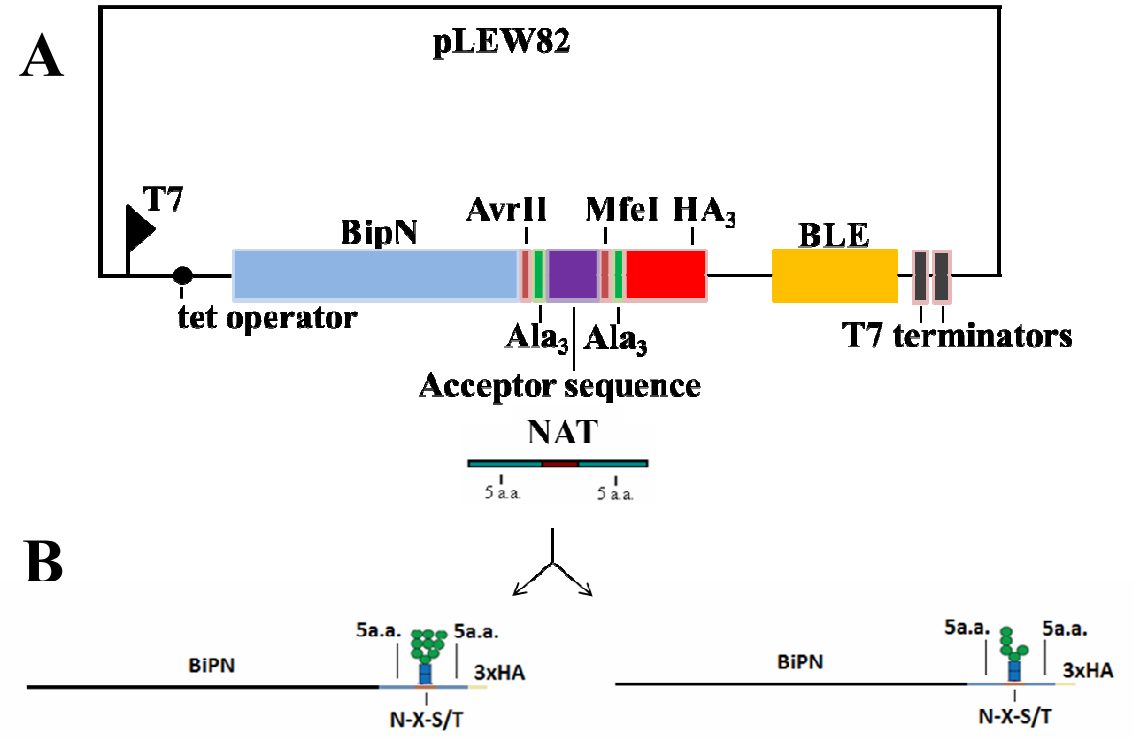


Figure 3.3.1. Glycosylation reporterconstruct. The pLEW82 plasmid drives expression of both the BipN fusion protein and the phleomycin resistance gene via a T7 promotor. The plasmid is tetracycline-inducible and has one tetracycline operator and two T7 terminators. *Tb*BipN (blue) is fused to anNAT sequon flanked by five variable residues (purple) followed by Ala linker regions (green), AvrII and MfeI restriction sites (pink) and an HA_3 epitope tag (red). When *Tb*BipN is expressed, the protein can carry a biantennary paucimannose or complex glycan transferred by *Tb*STT3A or a triantennary oligomannose glycan transferred by *Tb*STT3B. It is also possible that both types of glycan can be transferred to the asparagine in the N-X-S/T sequon.

After transfection of wild-type parasites, selected phleomycin resistant cells were lysed and the lysate was loaded on a SDS-PAGE gel and subsequently Western blotted using anti-HA (green channel) antibodies(Figure 3.3.2).

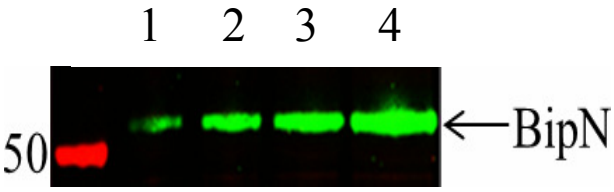


Figure 3.3.2. Western blot results after anti-Bipand anti-HA detection. Lane 1-4: Protein equivalent to 2, 4, 8 and 16*10^5 cells, respectively. Bands for *Tb*BipN were seen at approximately 55 kDa after anti-HA detection.

Protein equivalent to 2, 4, 8 and 16*10^5 cells were loaded in lane 1, 2, 3 and 4 and *Tb*BipN was detected at approximately 55 kDa in each lane.

3.3.2 Establishing an *in vivo* N-linked glycosylation assay

The first experimental sequon introduced was TEGLL**NAT**DEIAL. This sequence is found in VSG MITat.1.8 and is known to receive the biantennary $Man_5GlcNAc_2$ from *Tb*STT3A (Mehlert et al., 2010). The isoelectric point of the sequence is 3.42, thus fitting with the hypothesis that sequons in an acidic environment receive $Man_5GlcNAc_2$. To be able to visualise that the asparagine of the sequon had received a biantennary $Man_5GlcNAc_2$ from *Tb*STT3A, the cell lysate was treated with EndoH and PNGaseF before being loaded on a SDS-PAGE gel and Western blotted using anti-HA (section 2.8.2). Briefly, cells were lysed in 1 % SDS and 0.1 M DTT in the presence of protease inhibitors followed by dilution with Tx100 to 0.03 % SDS to facilitate adequate glycan digestion. The lysate was incubated overnight with either EndoH or PNGaseF. EndoH cleaves triantennary oligomannose glycans, leaving one GlcNAc residue still bound to the asparagine. PNGaseF is an amidase that cleaves between the innermost GlcNAc of

oligomannose, hybrid, and complex oligosaccharides and the asparagine residue from *N*-linked glycoproteins. During glycan removal, the modified asparagine residue is converted to aspartic acid. Hence, EndoH is only able to digest triantannary $Man_9GlcNAc_2$ that had been transferred by *Tb*STT3B, whereas PNGaseF can digest both $Man_9GlCNAc_2$ and biantennary $Man_5GlcNAc_2$ that had been transferred by *Tb*STT3A. Thus, a distinct digestion pattern, depending on what type(s) of glycan(s) are bound to the asparagine, can be visualised by Western blotting. The results after Western blotting are shown in Figure 3.3.3.

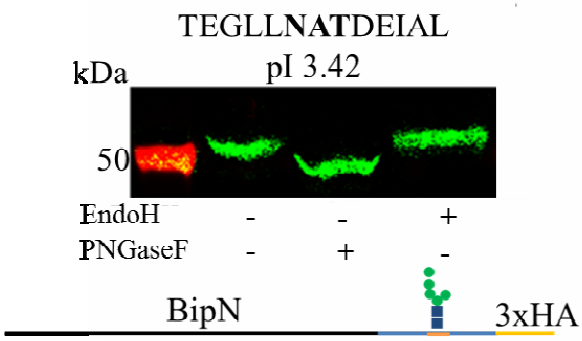


Figure 3.3.3. Western blot results after anti-HA detection. Molecular weight reduction is only seen following PNGaseF treatment. Thus, the asparagine in the TEGLL**NAT**DEIAL sequence had received an EndoH resistant biantennary $Man_5GlcNAc_2$ glycan from *Tb*STT3A. This result is in agreement with the hypothesis stating that sequons surrounded by acidic amino acids receive $Man_5GlcNAc_2$ from *Tb*STT3A (Izquierdo et al., 2009a).

It is evident that a molecular weight reduction due to glycan digestion, is only seen after treatment with PNGaseF. Since only PNGaseF was able to digest the glycan bound to

the asparagine in the TEGLL**NAT**DEIAL sequence, the glycan must be EndoH resistant and have originated from a biantennary $Man_5GlcNAc_2$, transferred by *Tb*STT3A (Figure 3.3.3). Hence, the *in vivo* assay, using *Tb*BipN as a glycosylation reporter protein, can distinguish between *Tb*STT3A and *Tb*STT3B transfer.

3.3.3 Further validation of the *in vivo* N-linked glycosylation assay

The *in vivo* assay indicated that it was possible to distinguish between *Tb*STT3A and *Tb*STT3B transfer,however, to be more confident that the assay was reliable, more positive controls were analysed using acceptor peptide sequences found in the *T. brucei* transferrin receptor. This protein is a dimer consisting of expression site associated gene 6 and of expression site associated gene 7 (ESAG6 and ESAG7) and has a total of eight glycosylation sequons(Mehlert et al., 2012; Steverding et al., 1994). Four of these have previously been experimentally characterised and the other four were predicted according to the pI values of the acceptor peptidesequences (Mehlert et al., 2012). Two of the eight sequon-containing sequences (ERNAL**NAT**AANKV and LEEMR**NAS**ALAAA) were found on both ESAG6 and ESAG7, therefore only six sequences were analysed by the *in vivo* assay. To generate *Tb*BipN constructs containing these acceptor peptides, forward and reverse DNA strands of the six sequences were synthesised individually in order to create duplex DNA with sticky ends upon annealing, as described in Figure 3.3.4. The annealed DNA sequences were subsequently cloned into the pLEW82 plasmid using the aforementioned AvrII and MfeI cloning sites.

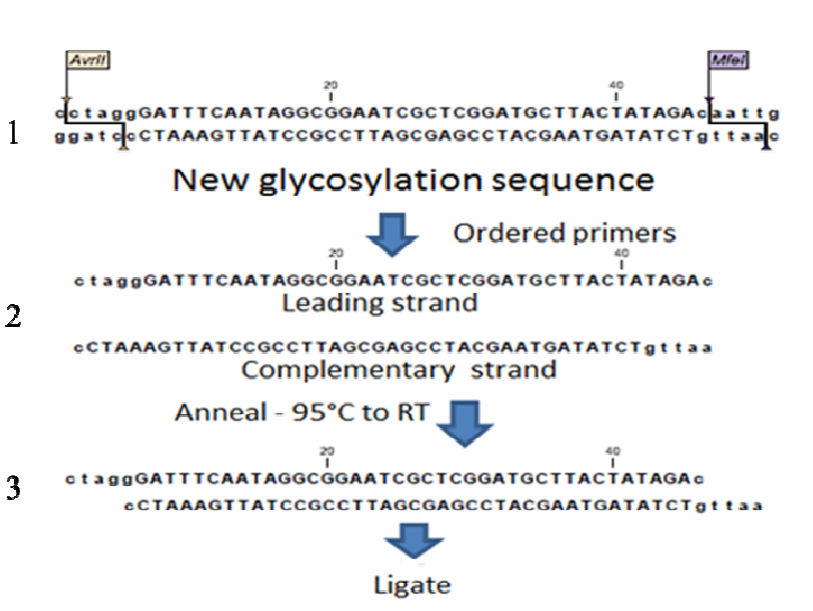


Figure 3.3.4. Strategy for engineering new N-glycosylation reporter constructs. Forward and reverse DNA sequences were synthesised individually and subsequently annealed together to create sticky ends without any digestion steps. The annealed DNA sequence was then cloned into the pLEW82 plasmid using AvrII and MfeI restriction enzyme sites.

The N-glycosylation of the six *Tb*BipN fusion proteins was analysed as described in section 3.3.2 and the Western blot results are shown in Figure 3.3.5.

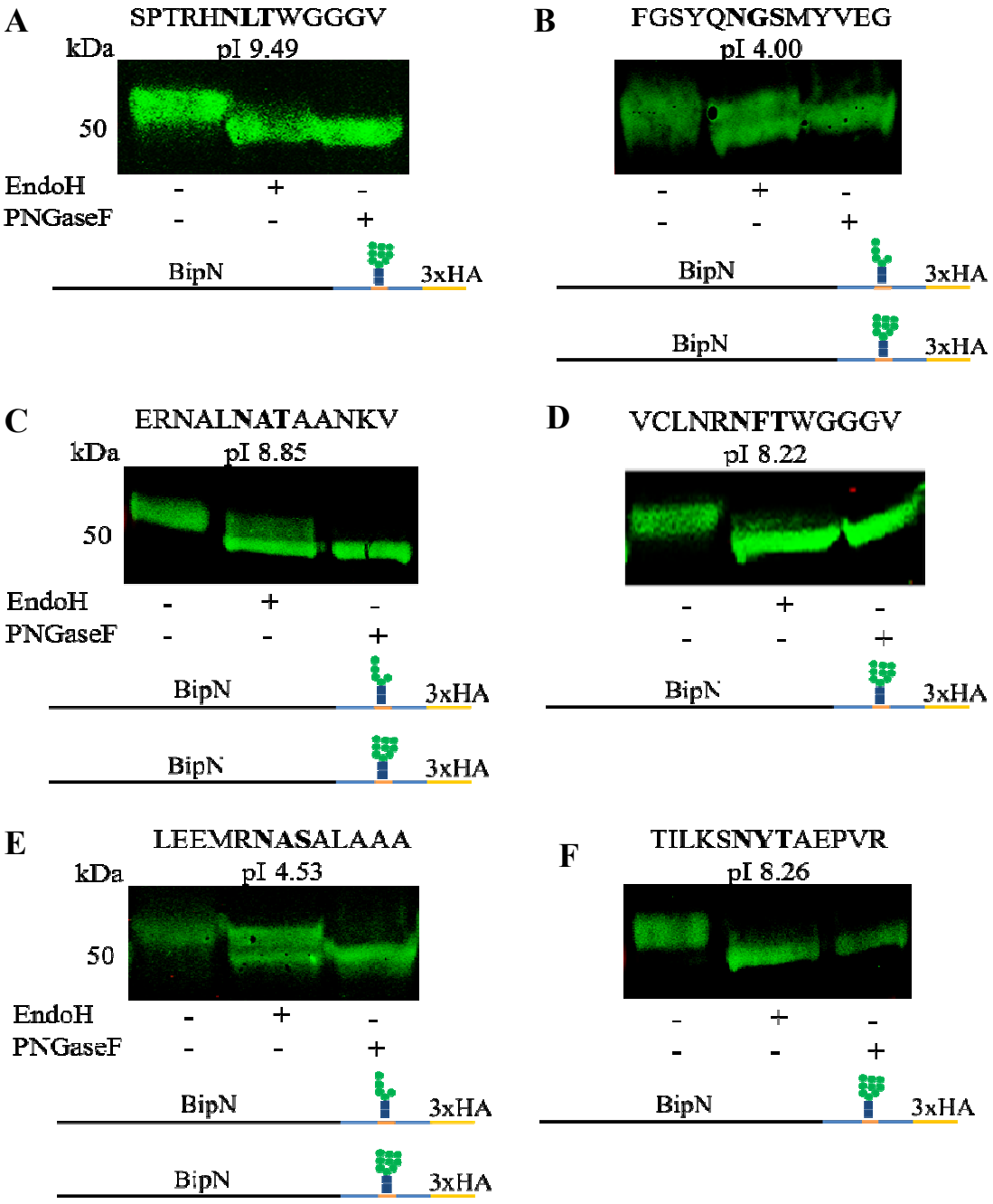


Figure 3.3.5. Western blot results after anti-HA detection. **A:** The glycans bound to the asparagine in the SPTRH**NLT**WGGGV sequence were sensitive to EndoH treatment, suggesting that $Man_9GlcNAc_2$ had been transferred by *Tb*STT3B. **B:** The glycans bound to the asparagine in the FGSYQ**NGS**MYVEG sequence were both resistant and sensitive to EndoH treatment, suggesting that both $Man_5GlcNAc_2$ and $Man_9GlcNAc_2$ had been transferred by *Tb*STT3A and *Tb*STT3B.

Figure 3.3.5.(continued) **C:** The glycans bound to the asparagine in the ERNAL**NAT**AANKV sequence were both resistant and sensitive to EndoH treatment, suggesting that both $Man_5GlcNAc_2$ and $Man_9GlcNAc_2$ had been transferred by *Tb*STT3A and *Tb*STT3B. However, $Man_9GlcNAc_2$ was the preferred glycan transferred to the sequon, since the lower band was much stronger. **D:** The glycans bound to the asparagine in the VCLNR**NFT**WGGGV sequence were sensitive to EndoH treatment, suggesting that $Man_9GlcNAc_2$ had been transferred by *Tb*STT3B. **E:** The glycans bound to the asparagine in the LEEMR**NAS**ALAAA sequence were both resistant and sensitive to EndoH treatment, suggesting that both $Man_5GlcNAc_2$ and $Man_9GlcNAc_2$ had been transferred by *Tb*STT3A and *Tb*STT3B. **F:**The glycans bound to the asparagine in the TILKS**NYT**AEPVR sequence were sensitive to EndoH treatment, suggesting that $Man_9GlcNAc_2$ had been transferred by *Tb*STT3B.

When comparing the results in Figure 3.3.5 with the previous experimental and predicted results (Mehlert et al., 2012), a positive correlation was seen. The same OST enzyme (*Tb*STT3A or *Tb*STT3B) was shown to be utilizing these acceptor peptides, however, the Western blot results often indicated overlapping sequon preferences for *Tb*STT3A and *Tb*STT3B, displayed by the double bands. For example, the sequons in ERNAL**NAT**AANKV and LEEMR**NAS**ALAAA were previously predicted to receive $Man_9GlcNAc_2$ and $Man_5GlcNAc_2$, respectively, but the two *Tb*BipN fusion proteins received both types of glycans. Although, the upper band in the ERNAL**NAT**AANKV Western blot was very weak, suggesting that most received oligosaccharides were the triantennary $Man_9GlcNAc_2$ transferred by *Tb*STT3B. A summary of the acceptor peptides investigated, their pI values and the experimentally (E) determined or predicted

(P) utilization by *Tb*STT3A and *Tb*STT3B from the Mehlert *et al.* study compared to the result of the *Tb*BipN *in vivo* study is seen in Table 3.3.1.

Table 3.3.1. List of TfR sequons, their pI values, the experimentally determined (E) or predicted (P) utilization by *Tb*STT3A and/or *Tb*STT3Band the results of this study, experimentally determined using the *in vivo* assay (adapted from Mehlert et al., 2013).

Site	**Sequon**	**pI**	**Transferring enzyme***	**Transferring enzyme****
ESAG6/7	ERNAL**NAT**AANKV	8.9	*Tb*STT3B (E)	*Tb*STT3A/B
ESAG6	LEEMR**NAS**ALAAA	4.5	*Tb*STT3A (E)	*Tb*STT3A/B
ESAG7	LEEMR**NAS**ALAAA	4.5	*Tb*STT3A (P)	*Tb*STT3A/B
ESAG6	TILKS**NYT**AEPVR	8.3	*Tb*STT3B (E)	*Tb*STT3B
ESAG6	SPTRH**NLT**WGGGV	9.4	*Tb*STT3B (P)	*Tb*STT3B
ESAG6	FGSYQ**NGS**MYVEG	4.0	*Tb*STT3A (P)	*Tb*STT3A/B
ESAG7	VCLNR**NFT**WGGGV	8.2	*Tb*STT3B (P)	*Tb*STT3B

*Mehlert et al., 2013

***Tb*BipN *in vivo* assay (this study)

3.3.4 N-linked glycosylation of *Tb*STT3A and *Tb*STT3B

With a robust *in vivo* assay, we wanted to investigate the sequon environments where the bound glycan type was unknown. Both *Tb*STT3A and *Tb*STT3B have N-X-S/T sequons, predicted to be found in loops on the protein surface (section 3.3.5), and are likely to be N-glycosylated. *Tb*STT3A has three sequons which are all shared with *Tb*STT3B, which has an additional fourth N-X-S/T site (Table 3.3.2).

Table 3.3.2 List of acceptor peptide sequences found in *Tb*STT3A and *Tb*STT3B.

Protein	Sequence	pI
*Tb*STT3A/B	ITGIG**NRT**TLADG	6.2
*Tb*STT3A/B	IYKVM**NVS**EESKA	6.5
*Tb*STT3A/B	YLWLR**NST**PEDAR	6.5
*Tb*STT3B	DFNRR**NRS**DAYYR	10.1

One of the sequences (DFNRR**NRS**DAYYR) found in *Tb*STT3B has an isoelectric point value of 10.1, suggesting that an EndoH sensitive $Man_9GlcNAc_2$ will be transferredto the asparagine in the sequon. The other sequons (found in both *Tb*STT3A and *Tb*STT3B) had pI values close to neutral, giving no indication of preferred glycan transfer. Following electroporation, recovered cells were lysed as described above (section 3.3.2) and the lysate was diluted and incubated with either PNGaseF or EndoH. Samples were loaded on a SDS-PAGE gel and Western blotted using anti-HA antibody. The results are seen in Figure 3.3.6.

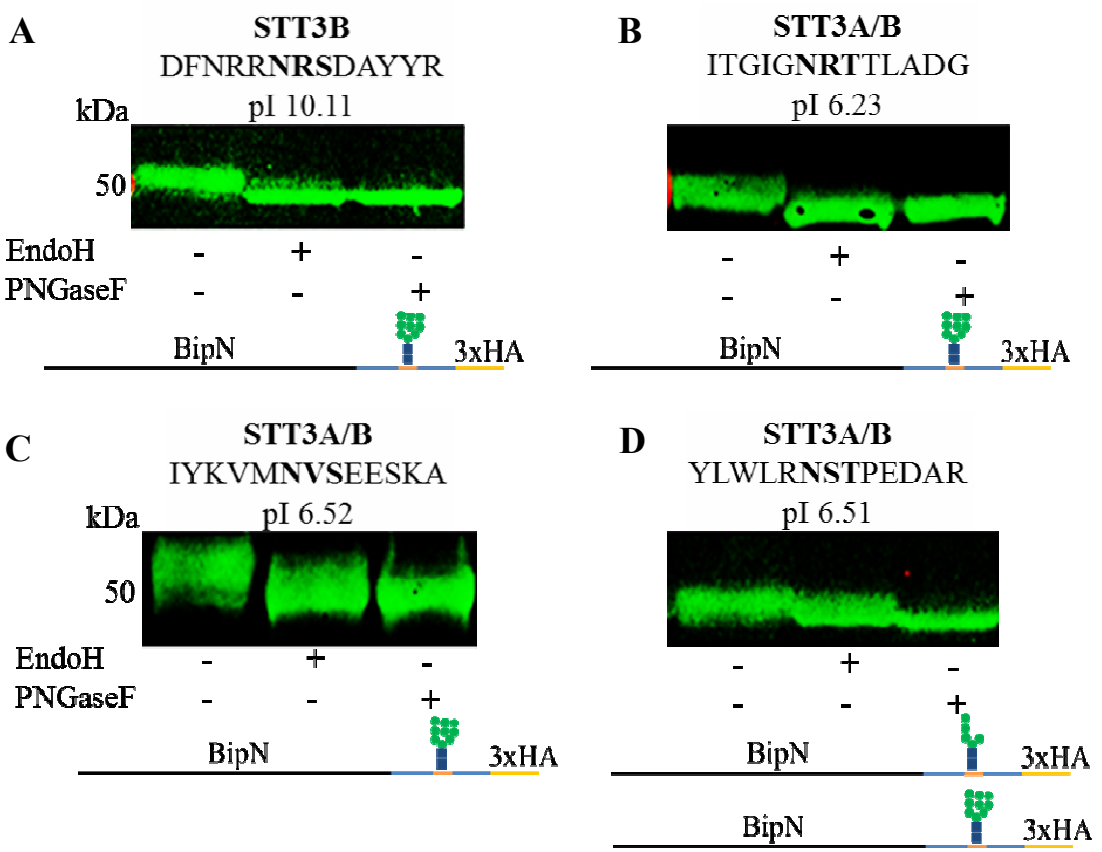


Figure 3.3.6. Western blot results after anti-HA detection. **A**, **B** and **C:** The glycans bound to the asparagines in the DFNRR**NRS**DAYYR, ITGIG**NRT**TLADG and IYKVM**NVS**EESKA sequences were EndoH sensitive, suggesting that a triantennary $Man_9GlcNAc_2$ had been transferred by *Tb*STT3B. **D:** Changes in mobility following both EndoH and PNGaseF treatment indicated that there were two types of glycans bound to the asparagine in the YLWLR**NST**PEDAR sequence, triantennary oligomannose oligosaccharides (EndoH sensitive) and biantennary oligosaccharides, hybrid and/or complex (EndoH resistant).

The asparagine within the DFNRR**NRS**DAYYR sequon peptide (*Tb*STT3B) was shown to receive an EndoH sensitive triantennary $Man_9GlcNAc_2$ from *Tb*STT3B, which was predicted from the relatively high pI value of 10.11 (Figure 3.39 A). A similar digestion pattern was seen for B and C in Figure 3.3.6, indicating that the asparagine within the

ITGIG**NRT**TLADG (pI 6.23) and IYKVM**NVS**EESKA(pI 6.52) sequon peptides, found in both *Tb*STT3A and *Tb*STT3B, had received an EndoH sensitive triantennary $Man_9GlcNAc_2$, transferred by *Tb*STT3B. The last peptide, YLWLR**NST**PEDAR (pI 6.51), is also found both in *Tb*STT3A and *Tb*STT3B and the digestion pattern was different compared to the results from the other sequon peptides. It was evident that the mobility changed following EndoH treatment, even though the top of the band was at similar molecular weight as the untreated sample.This indicates that both types of glycans were bound to the asparagine in the sequon(Figure 3.3.6 D).

In summary, the results suggested that *Tb*STT3A is able to autoglycosylate one of its three sequons (also found in *Tb*STT3B), whereas *Tb*STT3B seems to be able to glycosylate all sequons found in the two oligosaccharyltransferase enzymes. The sequon suggested to be acted upon by *Tb*STT3A (YLWLR**NST**PEDAR)has two acidic amino acids downstream of the sequon which might influence *Tb*STT3A transfer, supporting the hypothesis suggested by Izquierdo and colleagues (Izquierdo et al., 2009a). The DFNRR**NRS**DAYYR and ITGIG**NRT**TLADG peptides also have some acidic residues flanking the sequon, however, the arginine found between the asparagine and the serine/threorine in both peptides might have a negative effect on *Tb*STT3A catalysed transfer. Intriguingly, IYKVM**NVS**EESKA also has two acidic amino acids downstream of the sequon, similar to YLWLR**NST**PEDAR, but seems to receive only an EndoH sensitive triantennary $Man_9GlcNAc_2$, transferred by *Tb*STT3B. The results also support the suggestion that *Tb*STT3B has a broader acceptor peptide tolerance around the glycosylation sequon.

3.3.5 Investigation of *Tb*STT3A and *Tb*STT3B acceptor peptide specificity

It has been previously published (Izquierdo et al., 2009a) that *Tb*STT3A prefers to transfer $Man_5GlcNAc_2$ to sequons that are surrounded by acidic amino acids and the results from *Tb*STT3A and *Tb*STT3B glycosylation (3.3.4) supported this hypothesis. Furthermore, Izquierdo *et al* suggested that *Tb*STT3C also displayed a preference for acidic acceptor peptides (based on its expression in *S.cerevisiae* and subsequent analysis of the glycoproteome). To get a better understanding of why *Tb*STT3A and *Tb*STT3C seem to prefer acceptor peptides with acidic residues surrounding the sequon, the amino acid sequences of *Tb*STT3A, *Tb*STT3B and *Tb*STT3C were compared (Figure 3.3.7).*Tb*STT3A and *Tb*STT3C have more positively charged amino acids than*Tb*STT3B. The positions in the sequence alignment where both TbSTT3A and TbSTT3C have either a lysine or an arginine residue, but TbSTT3B does not, are highlighted with red boxes in Figure 3.3.7. At position 397, an arginine is found in *Tb*STTA and *Tb*STT3C, compared to the histidine found in *Tb*STT3B. Histidine is also a basic amino acid but the pKa value of 6.10 is considerably lower than the pKa value of arginine (12.48). Furthermore, the structure also differs between the bulkier imidazole ring of histidine and the long, side chain of arginine. In position 406 and 665, arginines are found in *Tb*STT3A and *Tb*STT3C, whereas a glycine is found in *Tb*STT3B. In position 558, the positively charged lysine and arginine are found in *Tb*STT3A and *Tb*STT3C, compared to a neutral glutamine residue found in *Tb*STT3B. Finally, position 567 is occupied by arginines in *Tb*STT3A and *Tb*STT3C, whereas a glutamine is found in *Tb*STT3B. These positively charged amino acids, specific for *Tb*STT3A and *Tb*STT3C, might play a role in explaining the preference of acidic acceptor peptide sequons seen for these two enzymes.

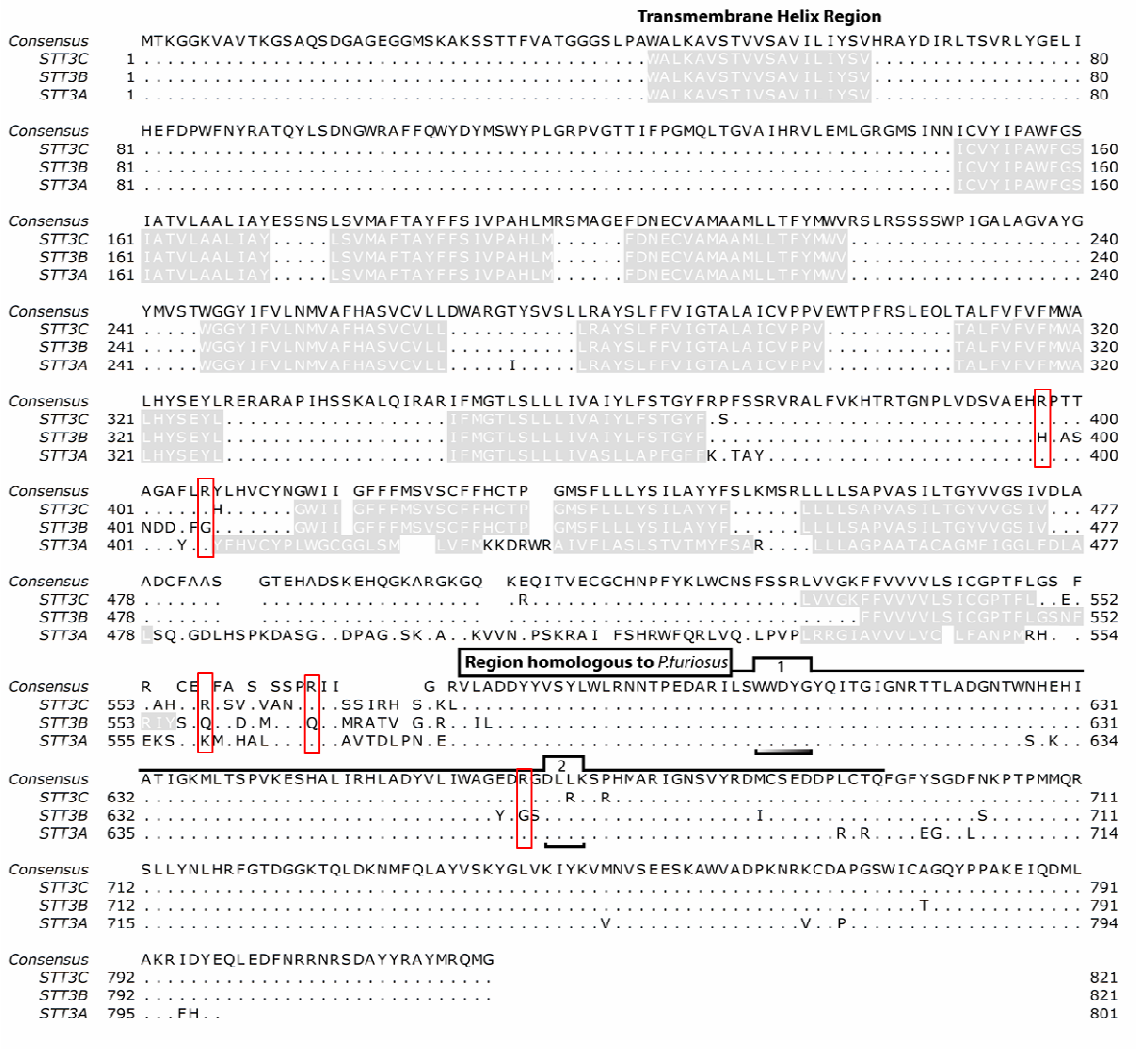


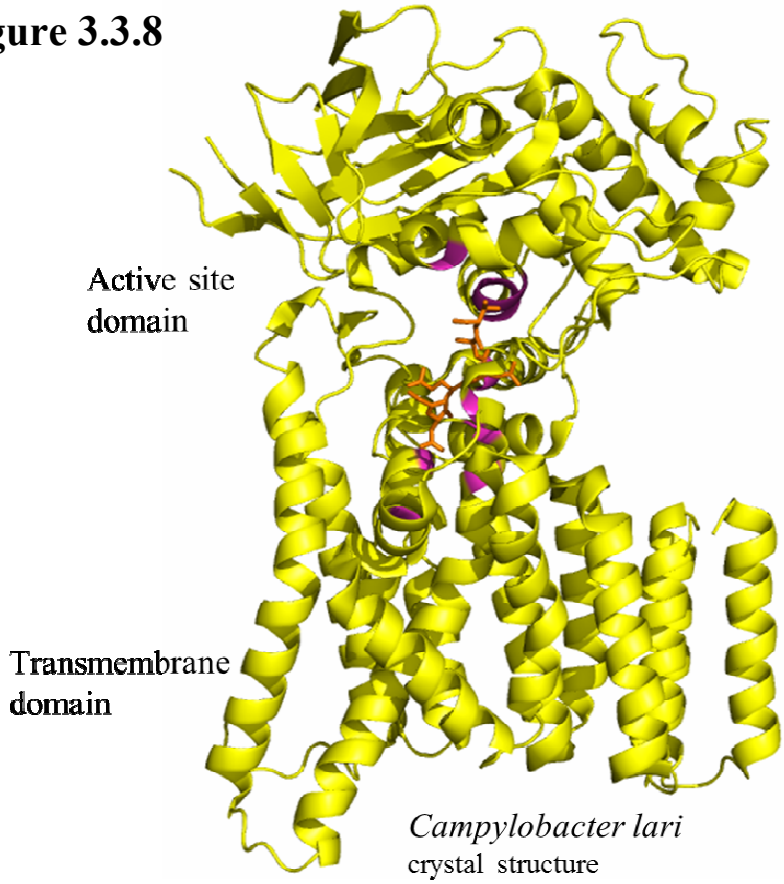
Figure 3.3.7. Protein sequence alignment of *Tb*STT3A, *Tb*STT3B and *Tb*STT3C. Positively charged amino acids found in *Tb*STT3A and *Tb*STT3C, not found in *Tb*STT3B, are highlighted in red boxes. Adapted from Izquierdo *et al*, 2009.

To this date, there are no structural data for the three oligosaccharyltransferase enzymes in *T. brucei*. Therefore, to be able to investigate if any of the positively charged residues specifically found in *Tb*STT3A and *Tb*STT3C are near the active site of the enzymes, models of the enzymes were built using Phyre2 (Kelley and Sternberg, 2009). Briefly, this program searches for homologues through a PSI-blast, then generates a hidden Markov model for a protein. At the same time it extracts the protein sequence for approximately 65.000 known structures and creates hidden Markov models for each of

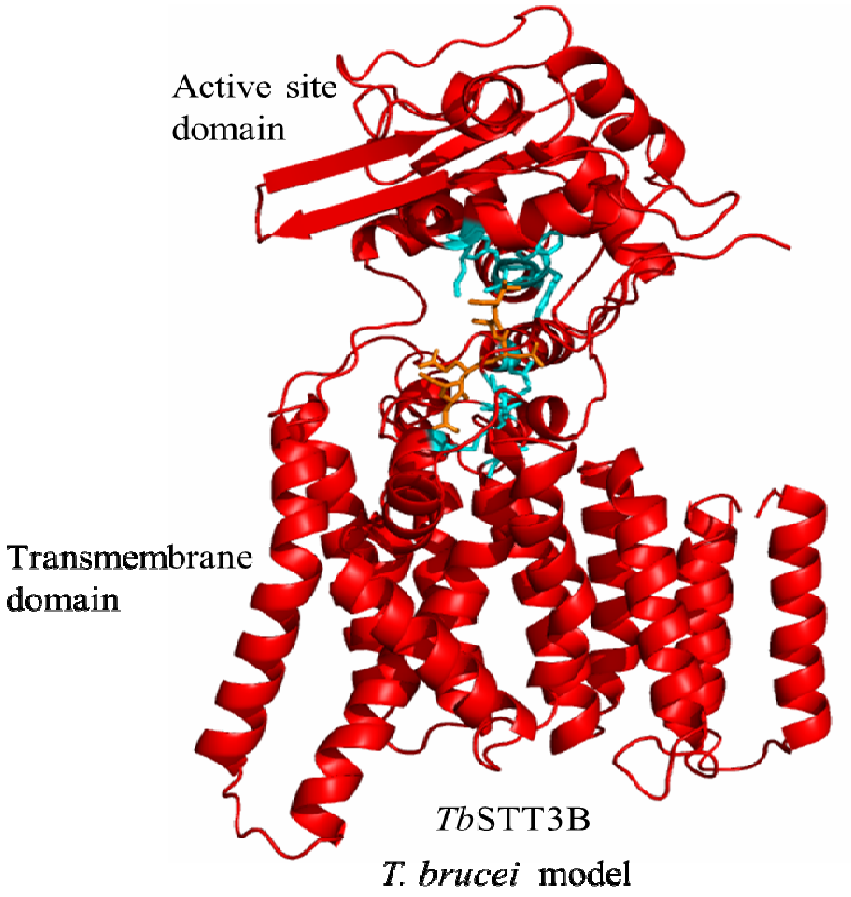
them, generating a database of these known models. Subsequently, the protein sequence is scanned through the abovementioned database to generate an alignment between the sequence being examined and the sequence of known structures. The top hit in this analysis was the crystal structure of *C. lari* OST (Lizak et al., 2011), followed by crystal structures of OST C-termini from various organisms. Although *C. lari* OST only share approximately 20 % of the protein sequence of *Tb*STT3A, *Tb*STT3B and *Tb*STT3C, 70 % of STT3A and 67 % of *Tb*STT3B and *Tb*STT3C protein sequences were modelled with 100 % confidence. Figure 3.3.8 showns the crystal structure of the PglB enzyme found in *C. lari* (yellow, A) and the model structures of *Tb*STT3B and *Tb*STT3C (red, B and purple, C) (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.).*Tb*STT3B and *Tb*STT3C have 95 % sequence identity (Izquierdo et al., 2009a)but still have different preferencesfor acceptor peptides. Therefore, these models are interesting to investigate, in trying to understand why *Tb*STT3C (and *Tb*STT3A) prefer sequons surrounded by acidic residues. The amino acids known to participate in catalysing the glycan transfer in *C. lari* are highlighted in magenta. These amino acids are conserved in *Tb*STT3B and *Tb*STT3C and were highlighted in cyan (*Tb*STT3B) and yellow (*Tb*STT3C). The acceptor peptide (GDQ**NAT**) that was bound to the *C. lari*oligosaccharyltransferase, is seen highlighted in orange (superimposed in *T. brucei* models).

Figure 3.3.8

A



B



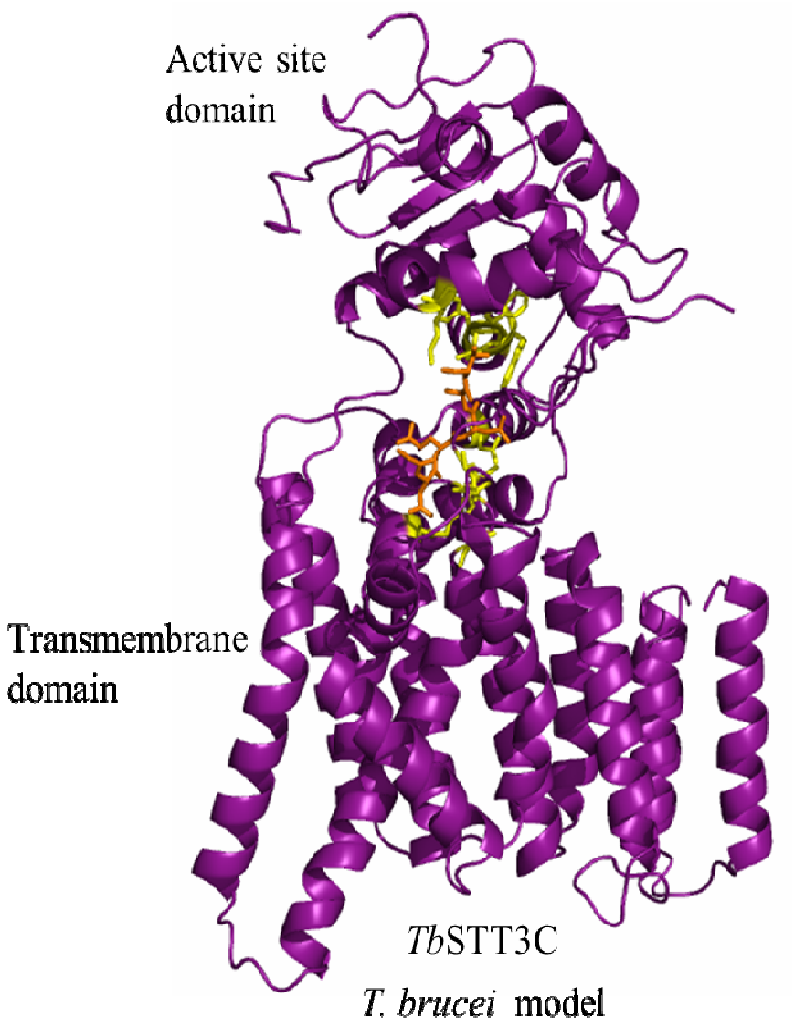
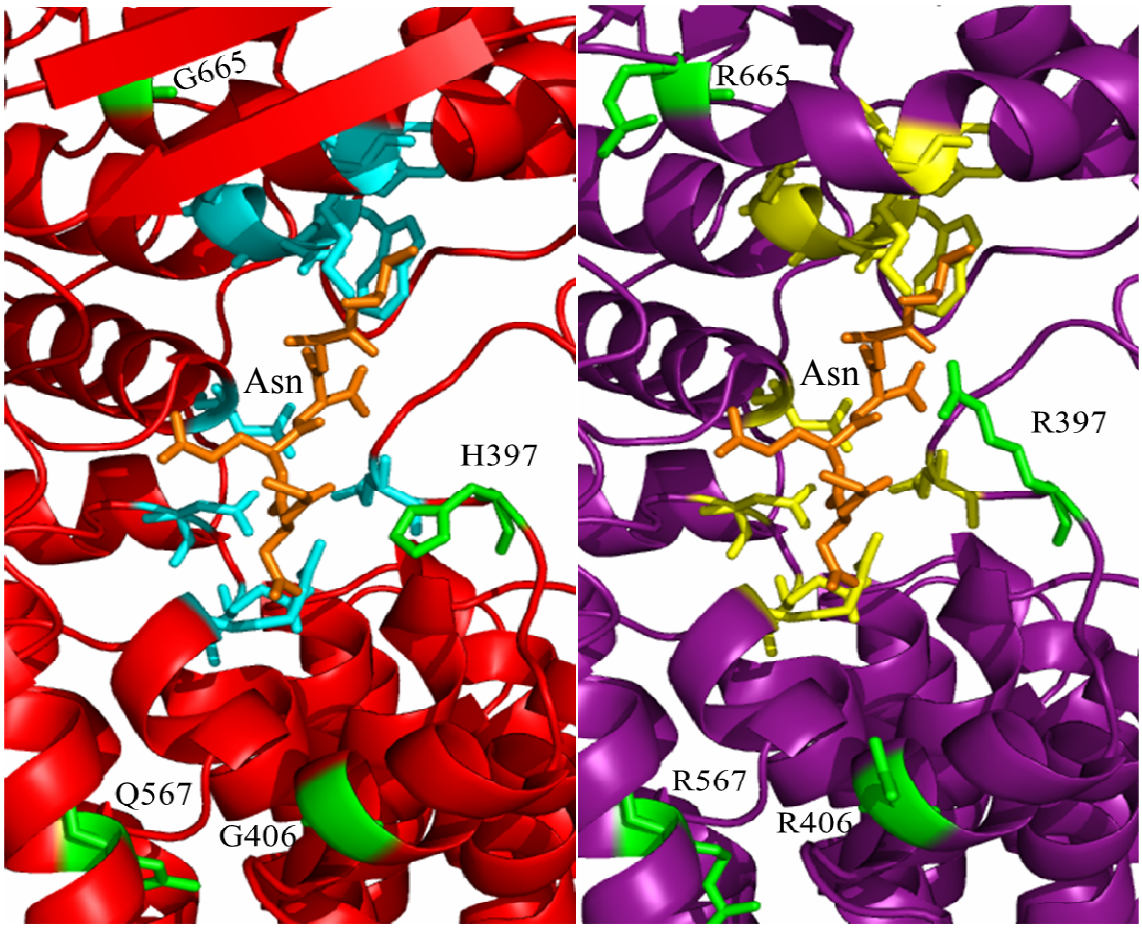


Figure 3.3.8. Crystal structure of *C. lari* and model structures of *Tb*STT3B and *Tb*STT3C.**A:** (Previous page) The *C. lari* oligosaccharyltransferase is seen in yellow, with amino acids proposed to be involved in glycan transfer highlighted in magenta. The bound acceptor peptide (GDQ**NAT**) is seen in orange. **B:** (Previous page) The model of *Tb*STT3B is seen in red, with conserved amino acids that presumably are involved in glycan transfer highlighted in cyan. The superimposed acceptor peptide (GDQ**NAT**) is seen in orange. **C:** The model of *Tb*STT3C is seen in purple, with conserved amino acids that presumably are involved in glycan transfer highlighted in yellow. The superimposed acceptor peptide (GDQ**NAT**) is seen in orange.

Four out of the five amino acid positions from Figure 3.3.9 are seen in the *Tb*STT3 models (residue 558 – K in *Tb*STT3A, Q in *Tb*STT3B and R in *Tb*STT3C – is not part of the models). All four of the observable amino acids are predicted to be positioned close to the active site and residue 397 (H in *Tb*STT3B and R in *Tb*STT3C) is positioned on the loop that has previously been reported to be involved in acceptor peptide binding in the *C. lari* oligosaccharyltransferase (Lizak et al., 2011). The four residues are seen in stick structure, highlighted in green. Furthermore, the superimposed acceptor peptide (GDQ**NAT**) that was bound to the *C. lari* oligosaccharyltransferase, is seen highlighted in orange and the conserved amino acids previously reported to take part in catalysing glycan transfer(Lizak et al., 2011)are highlighted in cyan (*Tb*STT3B) and yellow (*Tb*STT3C).

Figure 3.3.9.(Next page) Structural models of active sites of *Tb*STT3B (red) and *Tb*STT3C (purple). Amino acids that presumably are involved in glycan transfer highlighted in cyan (*Tb*STT3B) and yellow (*Tb*STT3C). The acceptor peptide (GDQ**NAT**) that bound to the *C. lari* oligosaccharyltransferase is superimposed and highlighted in orange. The residues which are positively charged in *Tb*STT3C but not in *Tb*STT3B are highlighted in green. The arginine at position 397 in *Tb*STT3C seems to point straight towards the acceptor peptide in the active site.



The arginine at position 397 points straight at the acceptor peptide and could perhaps be involved in selection of acidic glycosequon sequences in *Tb*STT3A and *Tb*STT3C.

Examining the differences in the amino acid sequences between *Tb*STT3B and *Tb*STT3C, suggests reasons why *Tb*STT3A and *Tb*STT3C prefer acceptor peptides with acidic residues surrounding the N-X-S/T sequon. Using the *in vivo* glycosylation assay it was possible to investigate how each position flanking the sequon affects *Tb*STT3A transfer. An aspartic acid scan was performed across the eleven Ala residues of the sequence shown in Table 3.3.3. All these peptides have the same pI value (3.10) and according to the current hypothesis they should all receive a biantennary $Man_5GlcNAc_2$ from *Tb*STT3A. A sequence with only alanine surrounding the sequon, with a pI of 6.01 and expected to receive $Man_9GlcNAc_2$ from *Tb*STT3B, was used as a negative control.

Table 3.3.3. List of sequences analysed by the *in vivo* glycosylation assay.

Name	Sequence	pI
Alanine control	AAAAA**NAT**AAAAA	6.01
D -5	DAAAA**NAT**AAAAA	3.10
D -4	ADAAA**NAT**AAAAA	3.10
D -3	AADAA**NAT**AAAAA	3.10
D -2	AAADA**NAT**AAAAA	3.10
D -1	AAAADNAT**AAAAA**	3.10
D +1	AAAAA**NDT**AAAAA	3.10
D +3	AAAAA**NAT**DAAAA	3.10
D +4	AAAAA**NAT**ADAAA	3.10
D +5	AAAAA**NAT**AADAA	3.10
D +6	AAAAA**NAT**AAADA	3.10
D +7	AAAAA**NAT**AAAAD	3.10

Following electroporation, recovered cells were lysed and the lysate was diluted and incubated with either PNGaseF or EndoH. Samples were loaded on a SDS-PAGE gel and Western blotted using anti-HA antibody. The SDS-PAGE gels were run over a longer time to separate potential double bands in order to facilitate band quantification using the Licor scanner software. The EndoH digestion indicates whether more than one glycan was transferred to the sequon by displaying two bands following Western blotting. If two bands are seen, the upper band indicates the *Tb*BipN molecules that received EndoH resistant $Man_5GlcNAc_2$ from *Tb*STT3A. The lower band indicates *Tb*BipN molecules which received EndoH sensitive $Man_9GlcNAc_2$ from *Tb*STT3B. Figure 3.3.10 shows the Western blot results after EndoH digestion of three of the transfected *Tb*BipN constructs.

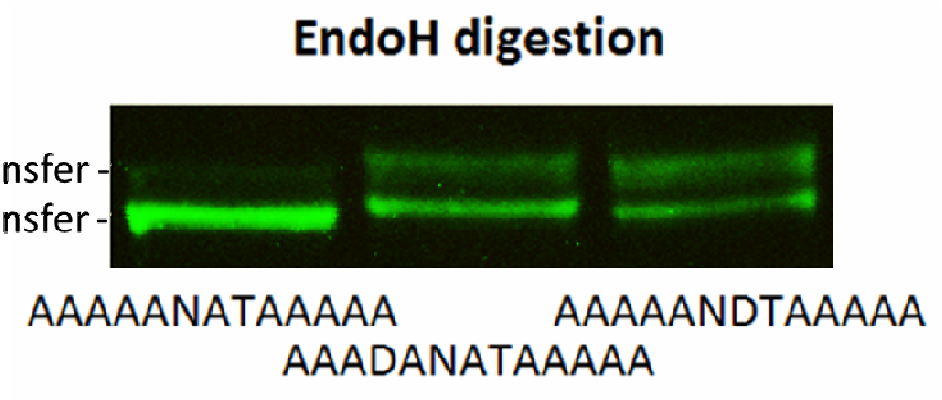


Figure 3.3.10. Western blot results after EndoH digestion. Double bands were seen in all lanes, suggesting that both *Tb*STT3A and *Tb*STT3B have transferred glycans to the sequons. The less intense band seen for the AAAAA**NAT**AAAAA sequence implies that *Tb*STT3A did not contribute much in the glycan transfer to the neutrally charged control sequence, in agreement with the current hypothesis .

There are double bands seen in all lanes in Figure 3.3.10, suggesting that both *Tb*STT3A and *Tb*STT3B transferred glycans to the sequons. However, the upper band of the AAAAA**NAT**AAAAA control sequence is weaker that in the other lanes, suggesting that *Tb*STT3A contributed less in glycan transfer. These bands could be quantified using the Licor scanner system software to calculate *Tb*STT3A transfer to each sequon-containing secuence. Three biological replicates, starting from the same cell density, were analysed three times each. PNGaseF and EndoH were added in excess to assure digestion going to completion. The collective results from all the analyses are seen in Figure 3.3.11.

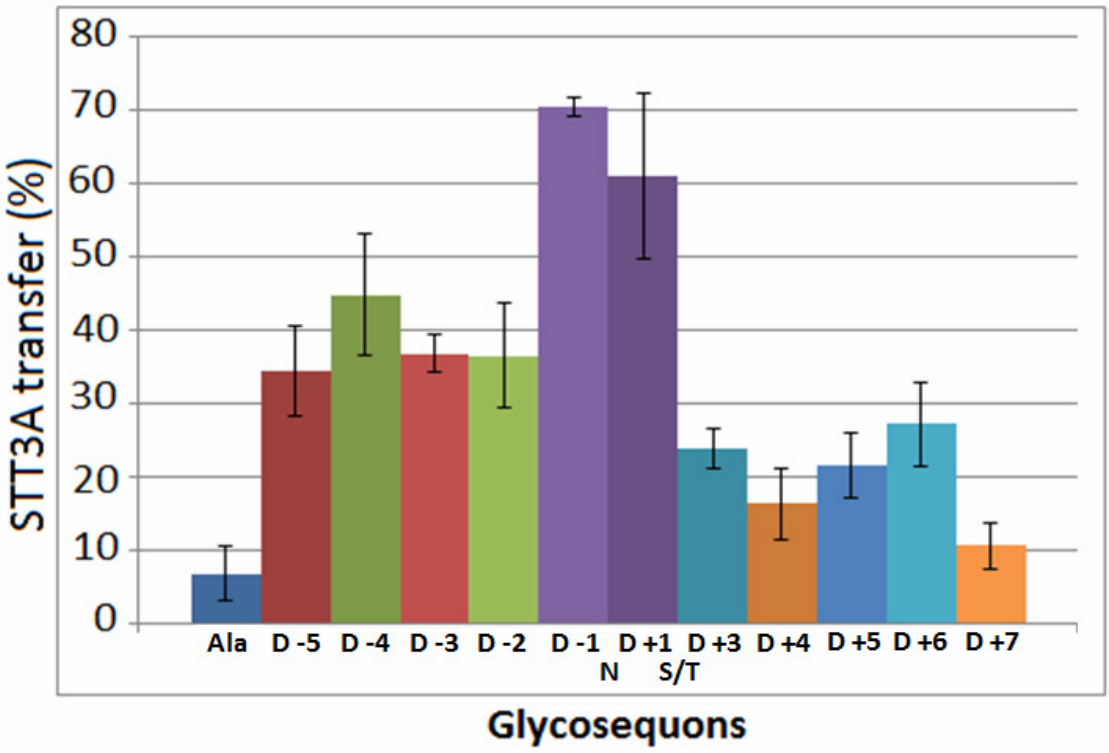


Figure 3.3.11. Graph visualising how *Tb*STT3A glycan transfer is affected by one single aspartic acid in and around the sequon. The negative control sequence (Ala) received approximately 5 % of its glycans from *Tb*STT3A. A single aspartic acid in any of the -5 to -2 upstream positions of the sequon generated approximately 35 % *Tb*STT3A transfer. A single aspartic acid in any of the +3 to +7 downstream positions influenced *Tb*STT3A transfer to between 10-25 %. The most striking effect, however, was seen when a single aspartic acid was introduced immediately adjacent to the asparagine of the sequon, which switched the *Tb*STT3A glycan transfer to 60-70 %.

The neutral control sequence, only containing alanines, showed about 5 % *Tb*STT3A transfer. This was expected since no acidic amino acids are found in the flanking positions of the sequon. Introducing a single aspartic acid into the sequon peptide in various positions reduces the pI uniformly and should allow more *Tb*STT3A transfer. However, although the overall pI of the region was uniformly reduced in all sequon

peptides, acidity in certain positions had a bigger impact on *Tb*STT3A transfer. An aspartic acid in the upstream positions -5, -4, -3 and -2 generates about 35 % *Tb*STT3A transfer, whereas the downstream positions 3,4,5,6 and 7 only ranges between 10-25 %. The most influential positions for *Tb*STT3A transfer were the sites immediately adjacent to the asparagine. A single aspartic acid in these positions switched *Tb*STT3A transfer to about 60-70 % compared to the 5% transfer when no aspartic acid was introduced. These results made us reassess the current hypothesis of acceptor peptide specificity.

This new hypothesis was supported by analysing sequon-containing sequences where the type(s) of glycan(s) attached to the asparagine where known. The data set (Rodriguez, Izquierdo, Procter and Ferguson, unpublished) was previously gathered following proteomic analysis subsequent to EndoH and PNGaseF digestion. Briefly,

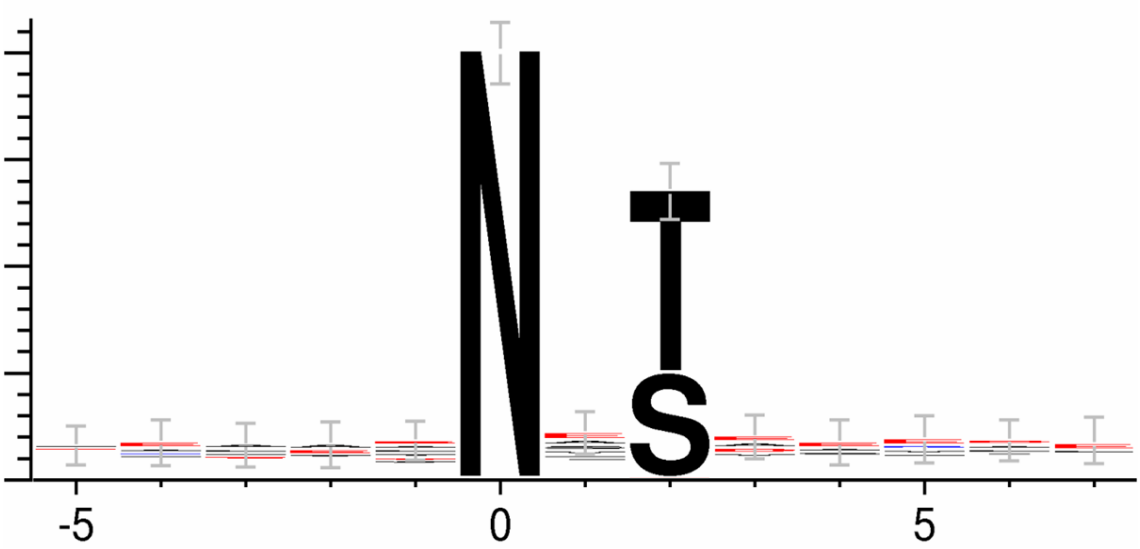
glycoproteins were captured from VSG-depleted cells followed by digestion with

EndoH and subsequently by PNGaseF. Hence, EndoH-sensitive glycopeptides appear

203 Da heavier by mass spectrometry due to the GlcNAc residue left attached to the

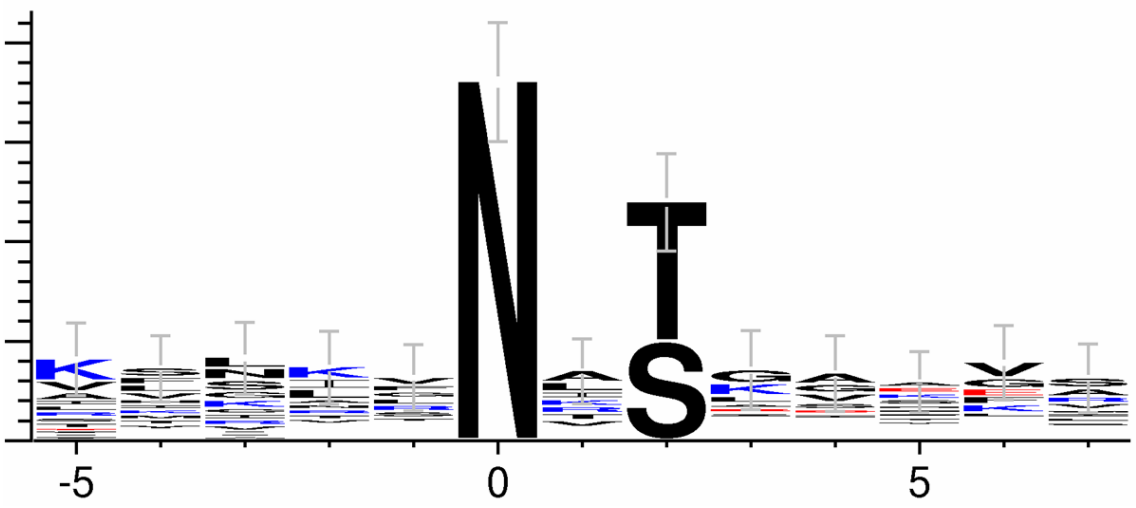
asparagine residues by endoH whereas the remaining endoH-resistant, but

PNGaseF-sensitive, sites appear 1 Da heavier because of the conversion of the asparagine residues to Asp by PNGaseF. The acceptor peptides that were shown to only receive a biantennary $Man_5GlcNAc_2$ from *Tb*STT3A (appendix) were used to create a motif logo (Figure 3.3.12). More acidic (red) than basic (blue) amino acids are found in these acceptor peptides, as expected. Also, more glutamic acids than aspartic acids were present, however, no other clear pattern was seen. Conversely, when looking closer at the sequon-containing sequences that had *Tb*STT3B transferred triantennary $Man_9GlcNAc_2$ bound to the asparagine, a pattern could be seen (Figure 3.3.13). Firstly, more basic (blue) than acidic (red) amino acids were found in the acceptor peptides, which was expected. Some acidic amino acids were seen, but they were mostly found downstream of the sequon. No acidic residues were seen in position -4, -3, -2 and +1 and only one acidic residue was seen in position -5 and -1. This is interesting as *Tb*STT3B acts upon the protein after co-translational glycosylation by *Tb*STT3A. Therefore, it seems like the acceptor peptides that contain a negatively charged amino acid in any of these positions are (mostly) acted upon by *Tb*STT3A, which is in agreement with the results seen in Figure 3.3.11.



Position	D	E	H	K	R
-5	7	9	0	7	4
-4	7	16	1	8	4
-3	7	8	1	3	6
-2	2	14	1	4	3
-1	7	12	1	1	6
0 (N)	0	0	0	0	0
+1	5	15	3	4	2
+2 (S/T)	0	0	0	0	0
+3	12	13	1	3	5
+4	5	16	0	1	4
+5	5	15	1	10	1
+6	12	5	4	6	3
+7	5	17	1	3	5

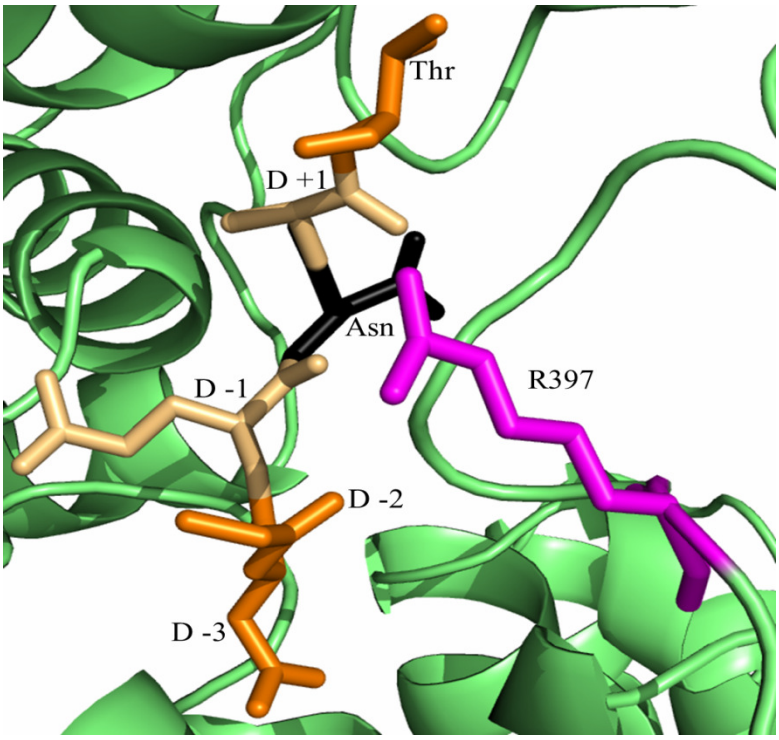
Figure 3.3.12. Motif logo of 90 acceptor peptides that had exclusively received biantennary $Man_5GlcNAc_2$ from *Tb*STT3A. The numbers of acidic (red box) and basic (blue box) amino acids in different positions of the acceptor peptides are highlighted. More negatively charged amino acids where found in and around the sequon compared to positively charged. Furthermore, more glutamic acids than aspartic acids were seen except in position +6.



Position	D	E	H	K	R
-5	1	0	0	8	2
-4	0	0	2	2	1
-3	0	0	1	3	2
-2	0	0	1	5	2
-1	0	1	1	1	3
0 (N)	0	0	0	0	0
+1	0	0	0	3	3
+2 (S/T)	0	0	0	0	0
+3	1	2	1	5	1
+4	2	1	1	1	1
+5	1	3	2	3	1
+6	1	3	0	3	1
+7	1	1	0	1	3

Figure 3.3.13. Motif logo of 29 acceptor peptides that had exclusively received triantennary $Man_9GlcNAc_2$ from *Tb*STT3B. The numbers of acidic (red box) and basic (blue box) amino acids in different positions of the acceptor peptides are highlighted. More positively charged amino acids (blue) where found in and around the sequon compared to negatively charged residues (red). Furthermore, no acidic residues were seen in position -4, -3, -2 and +1 and only one negatively charged residue was seen in position -5 and -1. This suggests that peptides with acidic residues in these upstream positions had already been acted upon by *Tb*STT3A.

The results seem to suggest that acidity in the upstream positions (-5, -4, -3, -2) of a glycosylation sequon are more important for selective *Tb*STT3A glycan transfer than the downstreams positions (+3, +4, +5, +6, +7). Interestingly, according to the results in Figure 3.3.11, a single aspartic acid in any of the immediately adjacent positions (-1 or +1) significantly improves*Tb*STT3A glycan transfer. Just like *Tb*STT3C (Figure 3.42), *Tb*STT3A also has an arginine at position 397 which seems to point straight into the active site where the acceptor peptide binds (Figure 3.3.14) This arginine is predicted to be found in a loop which has previously been reported to be a part of acceptor peptide binding in *C. lari* (Lizak et al., 2011, 2014).In Figure 3.3.14, the active site of *Tb*STT3A with a superimposed acceptor peptide (GDQ**NAT**) is seen. The asparagine of the sequon is highlighted in black, the two immediately adjacent positions (-1 and +1) in beige and the rest of the acceptor peptide in orange. The arginine in position 397 is highlighted in magenta.



*Tb*STT3A

Figure 3.3.14 Structural model of active site of *Tb*STT3A. An acceptor peptide (GDQ**NAT**) is superimposed and the asparagine of the glycosylation sequon is shown in black. The two immediately adjacent positions (-1 and +1) are beige and the rest of the peptide is orange. The asparagine in position 397 (magenta) seems to point at the acceptor peptide and may be involved in recognition and binding. This positively charged amino acid is found in a loop which has previously been reported to contribute to acceptor peptide binding in *C. lari*(Lizak et al., 2011, 2014).

Taken together, these data give a greater insight of *Tb*STT3A and *Tb*STT3B acceptor sequence specificity in *T. brucei*. Although the overall pI gives a good indication of selective recognition by *Tb*STT3A, we have shown that upstream amino acids and particularly amino acids immediately adjacent to the asparagine play a crucial role in *Tb*STT3A and *Tb*STT3B acceptor peptide specificity. Whether the effects seen in

Figure 3.3.11 are additive or whether the effect of the most significant acidic amino acids is dominant remains to be determined. Similarly, we need to assess whether glutamate has a similar, greater or lesser effect than aspartate. Furthermore, it is interesting that the arginine in position 397 in the *Tb*STT3A and *Tb*STT3C models is predicted to point towards the acceptor peptide, seemingly being close to the two immediately neighbouring residues of the asparagine. Additionally, the arginine is, according to the model, positioned in a loop that has previously been shown to facilitate binding of acceptor peptides in *C. lari*(Lizak et al., 2011, 2014). Whether this positively charged amino acid plays a role in recognising and binding acidic residues upstream and/or immediately adjacent the asparagine, remains to be determined.

4 – Discussion

4.1 Essentiality of *Tb*STTA - generation of *Tb*STT3A gene replacement mutants

In two separate attempts to create a *Tb*STT3 conditional null mutant, the data from the PCR screens and the Southern blot experiments (Figure 3.1.11, 3.1.13, 3.1.14 and 3.1.16) showed that BSD had successfully replaced the last endogenous *Tb*STT3A copy, but that new (weak) *Tb*STT3-probe hybridising bands, which were clone specific, were seen in the gene replacement cells. Furthermore, the glycosylation pattern of sVSG following gene replacement (Figure 3.1.18), showed that the *Tb*STT3A function was retained under non-permissive conditions. These results might suggest that the cells have rearranged their genome during *Tb*STT3A replacement, which would imply that the tetracycline-inducible ectopic copy could not complement for the loss of the endogenous *Tb*STT3A gene, suggesting that it is essential for cell viability *in vitro*. Interestingly, a recent study has reported that gene rearrangement had occurred in the *Tb*STT3 locus when cells were exposed to lethal concentrations of the mannose specific Hippeastrum Hybrid Agglutinin (HHA) (Castillo-Acosta et al., 2013). HHA preferentially binds to oligomannose glycans that originate from *Tb*STT3B transfer of $Man_9GlcNAc_2$ and, after establishing a resistant cell line, it was shown that the endogenous *Tb*STT3B was replaced by two novel chimeric genes, *Tb*STT3B/C/B and *Tb*STT3C/B/C. It was postulated that the cells had rearranged the *Tb*STT3 locus in order to express *Tb*STT3 proteins that do not transfer $Man_9GlcNAc_2$ efficiently, thus reducing the endocytosis of HHA and keeping the cells alive (Castillo-Acosta et al., 2013).

In order to test if something similar occurred during *Tb*STT3A replacement, it should be possible to extract gDNA from the gene replacement cells and sequence the *Tb*STT3 locus. Additionally, to elucidate why the ectopic copy of was not able to complement

gene replacement of the last endogenous copy of *Tb*STT3A, it would be interesting to sequence the ectopic copy of *Tb*STT3A incorporated into the ribosomal RNA spacer locus to check for errors or mutations. For future gene replacement trials, it would be beneficial to use a C-terminally tagged ectopic copy of *Tb*STT3A, which enables detection by Western blotting to verify that the gene is induced correctly following addition of tetracycline. The pLEW100v5 plasmid is similar to the original pLEW100, but a few useful differences could make that plasmid a better choice for introduction of the ectopic *Tb*STT3A copy. The rRNA promoter gives pLEW100v5 a higher expression level which could be important during future *Tb*STT3A gene replacement trials.

An alternative strategy to test the essentiality of *Tb*STT3A and *Tb*STT3B *in vitro* might be the removal of the whole last *Tb*STT3 locus from the [*Tb*STT3A/*Tb*STT3B/*Tb*STT3C]+/- heterozygote cell line after introduction of different ectopic copies. In this approach, we would predict that ectopic expression of *Tb*STT3A and *Tb*STT3B using pLEW100v5 should allow us to successfully replace the whole of the remaining endogenous TbSTT3 locus, in a similar way to that used to generate the heterozygote (Izquierdo et al., 2009a). However, if either *Tb*STT3A and/or *Tb*STT3B is/are essential for growth *in vitro*, then we will be unable to replace the remaining endogenous *Tb*STT3 locus (Figure 4.1).

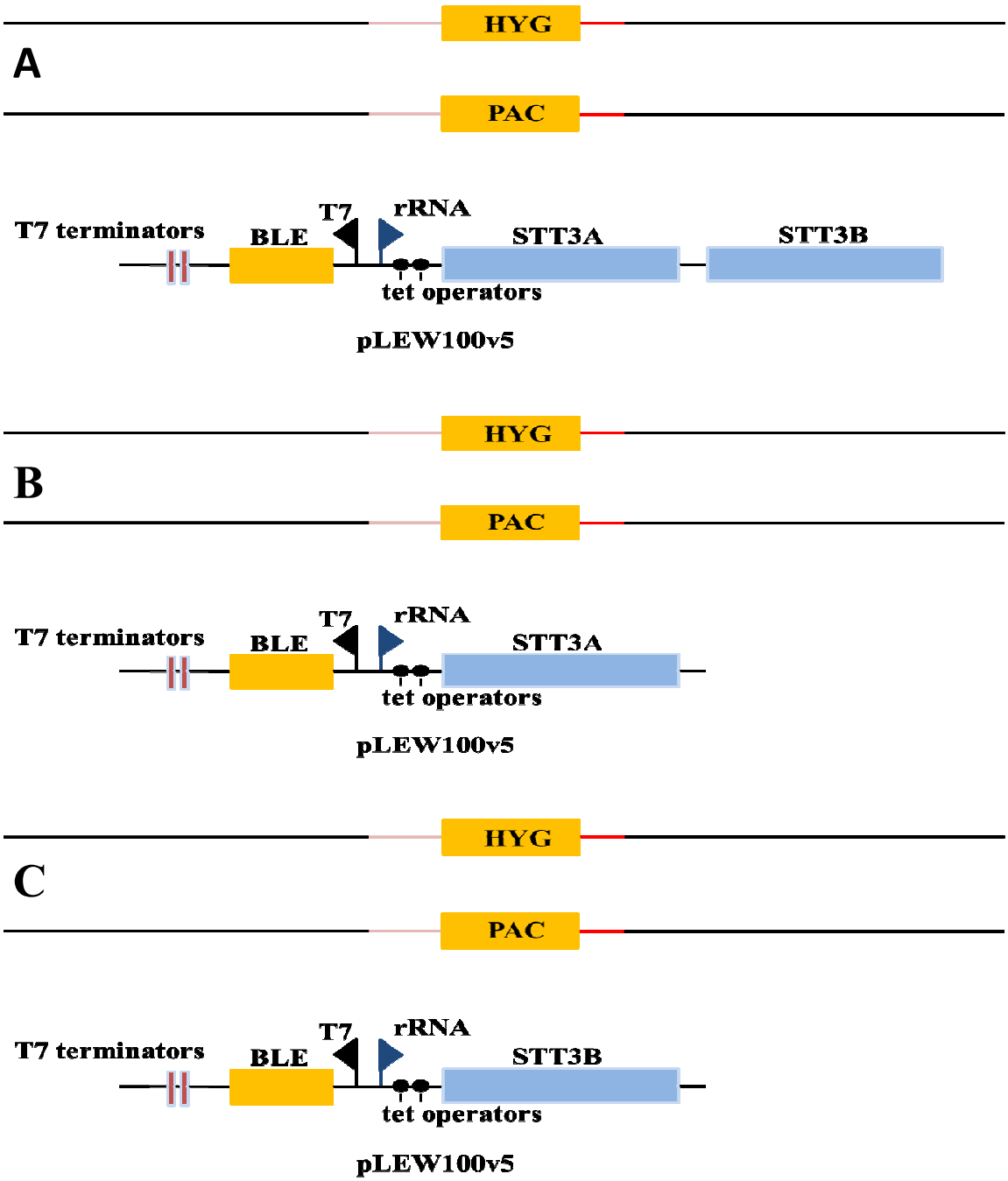


Figure 4.1. Alternative strategy to investigate *Tb*STT3A and *Tb*STT3B essentiality *in vitro.* Introduction of pLEW100v5 carrying a tetracycline-inducible ectopic copy of *Tb*STT3A and *Tb*STT3B (**A**), *Tb*STT3A (**B**) or *Tb*STT3B (**C**) followed by replacement of the last STT3 allele with HYG. Over expressing both *Tb*STT3A and *Tb*STT3B simultaneously should rescue allele replacement, while individual introduction of either would not support cell growth if the *Tb*STT3A and *Tb*STT3B genes are essential.

4.2 The search for novel *Tb*STT3A binding partners

Even though the *T. brucei* genome does not encode for any identifiable OST subunit homologues other than the three *Tb*STT3 genes(Berriman, 2005), we decided to investigate whether there were any novel OST subunits bound to *Tb*STT3A. This was important to do experimentally given the evolutionary distance of the kinetoplastids from other eukaryotic lineages. For example, previous studies in *T. brucei*have shown that novel clathrin-associating proteins were found and similarly, two novel subunits of the GPI transamidase werefound in *T. brucei* and shown to only have homologues in *Leishmania* and *Trypanosoma cruzi*(Adung'a et al., 2013; Nagamune et al., 2003).We therefore used *in situ* tagged *Tb*STT3A as bait to search for novel *Tb*OST subunits by blue native gel electrophoresis and immunoprecipitation experiments.

Before initiating immunoprecipitation under mild solubilisation conditions, optimisation of the percentage of digitonin was performed in order to minimise the risk of losing weakly associated partners, while still ensuring sufficient solubilisation of the membrane-bound *Tb*STT3A. A literature search suggested that 1 % digitonin was a common choice for both co-immunoprecipitation in *T. brucei* and for pulldowns of bigger complexes (for example OST) in other organisms (Oeljeklaus et al., 2012; Roboti and High, 2012a; Singha et al., 2012). However, it is clear in Figure 3.2.8, that, for example, the VSG dimer falls apart into the monomeric state when using 1 % digitonin. Further, since *Tb*STT3A solubilisation appeared to be efficient at 0.5 % digitonin we elected to use this detergent concentration in our further immunoprecipitation and blue native gels electrophoresis experiments.

The *Tb*STT3A-HA_3immunoprecipitationof SILAC labelled components gave relatively clear results concerning potential novel OST subunits in *T. brucei*. Both *Tb*STT3A (bait) and *Tb*STT3B were highly enriched in the pull-down and only three other proteins

had significantly enriched heavy to light isotope ratios (i.e., above the 3 sigma cut-off). However, these three proteins were 7-fold less enriched than *Tb*STT3B and are not known to localise in the ER, suggesting that *Tb*STT3B was the only protein specifically bound to *Tb*STT3A during the immunoprecipitation experiment. It is worth noticing that although *Tb*STT3A and potential associated proteins are expected to be specifically enriched, 500 other proteins were detected by mass spectrometry in the pull-down. Performing immunoprecipitation under SILAC methodology greatly simplifies recognition of true interacting proteins and can be used as a powerful tool for deconvoluting the complex datasets that typically emerge from pull-down experiments. The association between these two STT3 isoforms was further strengthened by the co-immunoprecipitation results (Figure 3.2.16). Taking these findings together with the large *Tb*STT3A-containing complex observed after blue native gel electrophoresis, we propose that *Tb*STT3A and *Tb*STT3B form a heteromeric OST complex in *T. brucei*. The reciprocal immunoprecipitation of *Tb*STT3B under SILAC conditions will be performed to validate the interactions between *Tb*STT3A and *Tb*STT3B and to confirm that no other potential OST subunit proteins are associated with *T. brucei* OSTs. Additionally, it would be tempting to use the *Tb*STT3A-HA_3–*Tb*STT3B-MYC_3 cell line to repeat the blue native gel experiment, enabling simultaneous detection of the two proteins in the high molecular weight complex.

Similar to *T. brucei*, the *Leishmania major* genome only encodes for the OST catalytic subunit STT3. Of the four STT3 genes in *Leishmania major*, three (*Tm*STT3A, *Lm*STT3B and *Lm*STT3D) were able to rescue STT3 deficiency in yeast cells. Further, the *Lm*STT3 proteins were not incorporated into the yeast OST complex (Nasab et al., 2008). Interestingly, these single subunit OSTs were suggested to form homodimers in their yeast host and*Lm*STT3A was even detected between approximately 500 and 700 kDa by Western blotting following blue native gel electrophoresis. However another

study proposed that *Lm*STT3A, *Lm*STT3B and *Lm*STT3D possessed anomalous migration behaviour and that they were functional as a monomeric OST in yeast (Hese et al., 2009). Despite these contradicting suggestions, it is interesting to question whether *L. major* STT3s can form homodimers, heterodimers or (as we speculate in *T. brucei*) multimeric OST complexes. Multiple copies of STT3 genes have also been reported in bacterial and archaeal genomes, opening up the possibility of more multimeric STT3 OST complex discoveries (Jervis et al., 2010; Magidovich and Eichler, 2009).

While higher eukaryotes have evolved increased complexity of OST subunit composition, which may reflect a diversity of protein substrates and sequons, *T. brucei*OST seems to consist solely of*Tb*STT3A and *Tb*STT3B, bound together, in bloodstream form parasites. Since the glycosylation efficiency has been suggested to be rather high in the parasite cells(Izquierdo et al., 2009a), it is interesting to question whether*T. brucei* OST has the glycosylation "fine tuning" capabilities described for the OSTs of higher eukaryotes. For example, the mammalian OST subunits MagT1/IAP and TUSC3/N33, and their Ost3 and Ost6 homologues in yeast, share lumenal thioredoxin domains(Cherepanova et al., 2014; Fetrow et al., 2001; Kelleher and Gilmore, 2006; Mohorko et al., 2014). It has been suggested that MagT1/IAP and TUSC3/N33 can form transient mixed disulfides with a free thiols in nascent glycoproteins, hence delaying disulfide bond formation and giving more time for glycosylation to occur. Furthermore, MagT1/IAP has been proposed to react with a disulfide in a nascent glycoprotein when in its reduced form, thus allowing access of STT3B to previously inaccessible sequons (Cherepanova et al., 2014). *T. brucei* STT3s do not have the CxxC motif found in MagT1/IAP and TUSC3/N33 and the Ost3 and Ost6, but a CxC site is present in *Tb*STT3B and *Tb*STT3C and this motif has been reported to have disulfide isomerase activity(Woycechowsky and Raines, 2003).

Interestingly, *Tb*STT3B has been shown to have a broader acceptor peptide preference than *Tb*STT3A (Izquierdo et al., 2009a) and perhaps this cysteine motif could add oxidoreductase features that expand the range of acceptable sequon environments forglycosylation by*Tb*STT3B. The CxC motif of *Tb*STT3B is predicted to be in the EL5 loop that has been shown to interact with the acceptor polypeptide (Lizak et al., 2011, 2014). Perhaps the cysteines in the CxC motif can form transient mixed disulphides with the nascent protein in order to enhance the efficiency of *Tb*STT3B glycosylation.

It would be possible to investigate the influence of cysteine residues adjacent to glycosylation sequons using the *Tb*BipN *in vivo* assay system described in this thesis. Thus, we could investigate whether the presence of adjacent Cys residues might reduce their recognition by *Tb*STT3A and increase their utilisation by *Tb*STT3B.

4.3 Acceptor peptide substrate specificity of *Tb*STT3A and *Tb*STT3B

The result of a single aspartic acid introduced into variable flanking positions clearly showed that a low pI value alone for acceptor peptides is not sufficient to explain the differences between *Tb*STT3A and *Tb*STT3B sequon specificity. Although the overall pI of the acceptor peptide was uniformly reduced by aspartic acid in all sequon peptides, acidity in certain positions had a bigger impact on *Tb*STT3A transfer and these results made us reassess the current hypothesis of acceptor peptide specificity. We showed that upstream amino acids and particularly amino acids immediately adjacent to the asparagine, play a dominant role in *Tb*STT3A and *Tb*STT3B acceptor peptide specificity. The *Tb*STT3B-utilised glycosylation site motif logo supported this view and, together with the molecular model data for the*T. brucei* STT3s, it is tempting to suggest a possible role for the arginine residue in position 397 (R397), found in *Tb*STT3A and *Tb*STT3C but not in *Tb*STT3B, playing a role in acceptor site

recognition by *Tb*STT3A and *Tb*STT3C. Since R397 is predicted to be in the C-terminal part of the flexible EL5 loop, which has been proposed to be involved in acceptor peptide binding(Lizak et al., 2011, 2014), it is possible that the basic arginine side-chain can 'scan' the sequon flanking residues and interact electrostaticallywith acidic amino acids and thereby enhance the recognition of specific sequons. It is important to be careful when making assumptions based on molecular models, however, more evidence about the involvement of R397 in *T. brucei* STT3 sequon specificity was published recently (Castillo-Acosta et al., 2013). The chimeric *Tb*STT3B/C/B protein described in this paper showed much less efficient recognition of the asparagine 428 (N428) glycosylation site in VSG221, normally the native choice of *Tb*STT3B. In other words, the *Tb*STT3B/C/B chimeric OST appeared to have attained a peptide acceptor specificity more similar to *Tb*STT3A than *Tb*STT3B. Within the C component of the *Tb*STT3B/C/B chimeric protein, there are only 11 amino acids changes. Interestingly, one of these is H397R. This supports the notion that the presence of R397 in *Tb*STT3A (and the *Tb*STT3B/C/B chimera) is the reason why these OSTs skip the N428 glycosylation site of VSG221 thathas no acidic residues near the sequon (KTGNT**NTT**GSSNS). Based on these observations, we have engineered point mutations at position 397 in *Tb*STT3B and *Tb*STT3C (initially H397R in *Tb*STT3B and R397H, R397D, R397Q in *Tb*STT3C) and these, together with wild type *Tb*STT3B and *Tb*STT3C, will be used to complement yeast OST mutants and analyse glycosylation site utilisation in yeast using an MS based method in Professor Markus Aebi's lab at ETH in Zurich. According to our current thinking, the H397R mutation in *Tb*STT3B should make this OST more like *Tb*STT3A and *Tb*STT3C with respect to acceptor peptide specificity, whereas R397H, R397D, R397Q in *Tb*STT3C might be expected to make this OST more like *Tb*STT3B with respect to acceptor peptide specificity.

Lastly, since charge seems to play an important role in *Tb*STT3 sequon specificity, it is noticeable that three of the other amino acids that changed in the *Tb*STT3B/C/B chimer were D402G, D403A and G406R. All these amino acid changes could be argued to reduce the ability of the enzyme to recognise the VSG221 N428 glycosylation site (KTGNT**NTT**GSSNS), when looking at their predicted locations in *Tb*STT3B (Figure 4.2).

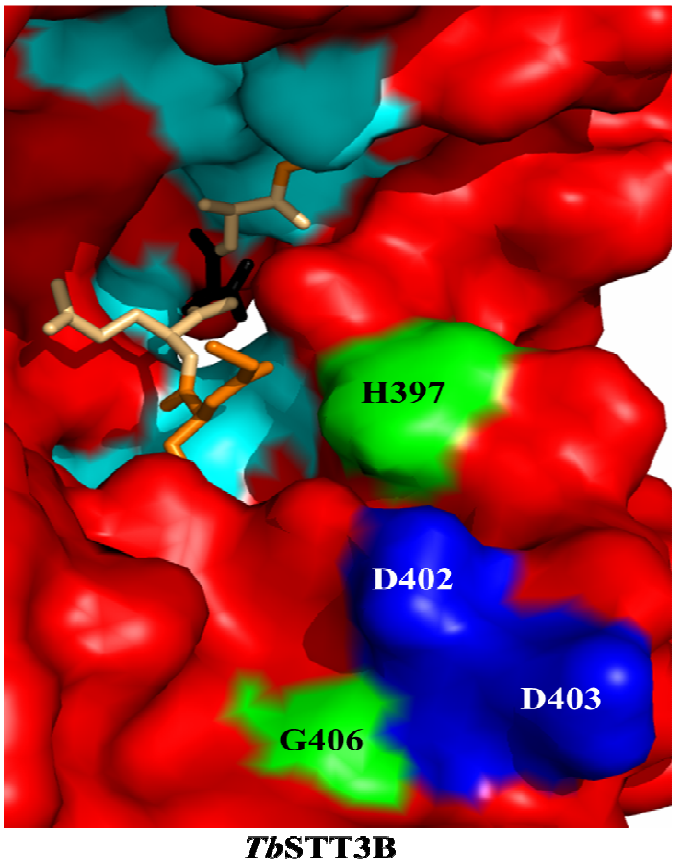


Figure 4.2. Model of the *Tb*STT3B active site. The superimposed acceptor peptide (GDQ**NAT**)is seen binding conserved residues that where involved in catalysing oligosaccharide transfer in *C. lari* (cyan) and the four residues involved in charge change in *Tb*STT3B/C/B are highlighted in green (changed into basic arginines) and blue (acidic residues changed into neutral).

In this model, the residues 397, 402, 403 and 406 line up just below the glycine in the superimposed acceptor peptide (GDQ**NAT**) and it is possible that a longer acceptor peptide could interact with these residues. Therefore, it is appealing to suggest that the lysine in the -5 position in the N428 acceptor peptide might interact with one of the two aspartic acids in *Tb*STT3B and that this interaction is lost in the *Tb*STT3B/C/B chimera, thus reducing recognition of the sequon. It is also interesting to note that the N263 sequon of VSG221, which is acted upon exclusively by *Tb*STT3A, has a glutamic acid immediately adjacent to the asparagine (TAEFR**NET**AGIAG), in agreement with the results showing that acidity in position -1 and +1 heavily influence *Tb*STT3A transfer of $Man_5GlcNAc_2$.

It might be possible to investigate how some of the residue changes seen in the *Tb*STT3B/C/B chimera (for example H397R, D402G and D403A) influence *Tb*STT3A and *Tb*STT3B glycosylation *in vivo* by using the protocol for *in situ* tagging *Tb*STT3A and *Tb*STT3B (Figure 4.3). In this way, different point mutations could be introduced into *Tb*STT3A and *Tb*STT3B that might switch the acceptor peptide specificities of the two proteins, such that *Tb*STT3A transfers $Man_5GlcNAc_2$ to a broad range of sequons while *Tb*STT3B transfers $Man_9GlcNAc_2$ to a reduced number of sequons, most likely those surrounded by acidic residues. This could subsequently be further tested by introducing the previously investigated (single aspartic acid) *Tb*BipN acceptor peptides to probe specificity in a systematic way.

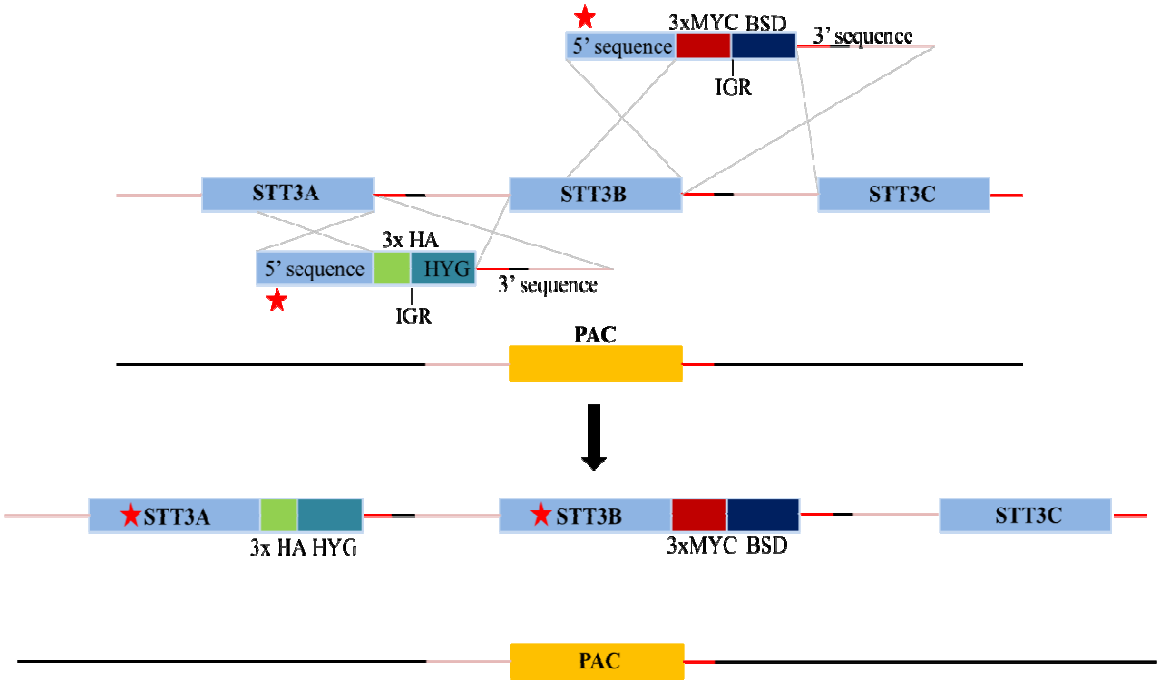


Figure 4.3. Strategy for mutation of the endogenous *Tb*STT3A and *Tb*STT3B using the *in situ* tagging protocol. Since a long stretch of the ORFs is used when *in situ* tagging, different mutations can be included in the tagging construct which will be transferred to the last allele of the endogenous genes via homologues recombination. The red star marks the point mutations.

5 – Appendix

5.1 pMOTag4M* construct for *in situ* tagging *Tb*STT3B

gctgggtaccgggccGCAGATTGTTTCGCCGCTTCCGGAACAGAGCACGCCGACAG
TAAGGAGCATCAAGGGAAAGCCCGTGGTAAGGGACAGAAGGAACAAATCA
CTGTCGAGTGTGGGTGCCATAATCCCTTCTACAAATTATGGTGCAATTCATT
TTCCTCCCGCCTGGTAGTTGGTAAGTTTTTTGTCGTTGTTGTCCTTGCCATCT
GTGGACCCACATTTCTTGGGTCTAACCTCCGGATATATTCTGAGCAATTCGC
AGACAGCATGTCGAGCCCCCAGATTATAATGAGGGCAACTGTCGGTGGACG
ACGAGTTATCTTGGATGATTACTACGTGTCGTACTTGTGGCTGCGAAACAAT
ACGCCTGAAGATGCCCGTATTCTCTCATGGTGGGACTACGGGTATCAAATCA
CTGGAATTGGCAATCGCACAACCCTTGCGGATGGTAACACATGGAATCACG
AGCACATAGCAACTATTGGAAAGATGCTTACATCCCCTGTGAAGGAGTCAC
ATGCTCTTATACGCCATCTCGCTGATTATGTGCTGATATGGGCCGGTTATGA
TGGCAGCGATTTACTTAAATCGCCACACATGGCTCGGATAGGCAACAGTGT
ATATCGCGATATATGCTCAGAGGATGATCCGCTGTGTACGCAGTTCGGGTTT
TATAGTGGTGACTTCAGTAAACCTACGCCTATGATGCAGCGGTCCCTATTAT
ACAATCTGCACAGGTTTGGTACGGATGGCGGGAAGACACAACTGGATAAGA
ACATGTTTCAGCTCGCCTACGTGTCAAAGTATGGTTTGGTGAAGATCTACAA
GGTGATGAATGTGAGTGAAGAGAGCAAGGCGTGGGTTGCAGACCCAAAGA
ACCGCGTATGCGACCCACCCGGATCTTGGATATGCGCCGGCCAGTACCCGC
CAGCGAAGGAGATCCAAGACATGTTAGCGAAGAGGATTGACTACGAACAA
CTCGAGGATTTCAATCGCCGCAATCGAAGTGACGCTTATTATCGTGCGTATA
TGCGTCAGATGGGTCTGCAGGCGGGCGGCCGATCGAGATCCGAGGAGCAGA
AGCTGATCTCGGAGGAGGATCTGCTGAGATCCGAGGAGCAGAAGCTGATCT
CGGAGGAGGATCTGCTGAGATCCGAGGAGCAGAAGCTGATCTCGGAGGAG

GATCTGCTGTGATCTAGTGCGGCCGTCGACAAAGTGTGACAACGTCGCACC

ATGTGTAGGTTTTCATTTATGTTCTTTCTTTCTTTCTTTTGTGAATTTGTTTTC

TGTCTCAAATGTTTTTTAATTCGCTTGGGACCTATGTTTTTCTTGTTTTTTTGCT

CACCCTTTGTGTAGGAGGCACCCTGTCACGTCTGTGGTTGCGTGTATGCCTT

CCTTCCCCTTATTCGCTTCTTCCTGTCGTGTCACACCTCTTTCTCCCTCTCCCT

TTCCGCCTTTTCTTTCAATCTTGTTTTCTCGACCAGCCCTACTAGAGGAGAAA

GAATAGTAACCCTTTCATCAAAGAAAATAGTTCAAACGAATTGCTAGCTTA

AGCTTGATCTAGAACTAGTGATGGCCAAGCCTTTGTCTCAAGAAGAATCCA

CCCTCATTGAAAGAGCAACGGCTACAATCAACAGCATCCCCATCTCTGAAG

ACTACAGCGTCGCCAGCGCAGCTCTCTCTAGCGACGGCCGCATCTTCACTGG

TGTCAATGTATATCATTTTACTGGGGGACCTTGTGCAGAACTCGTGGTGCTG

GGCACTGCTGCTGCTGCGGCAGCTGGCAACCTGACTTGTATCGTCGCGATCG

GAAATGAGAACAGGGGCATCTTGAGCCCCTGCGGACGGTGCCGACAGGTGC

TTCTCGATCTGCATCCTGGGATCAAAGCCATAGTGAAGGACAGTGATGGAC

AGCCGACGGCAGTTGGGATTCGTGAATTGCTGCCCTCTGGTTATGTGTGGGA

GGGCTGACCCTGTGAAGATAAATTCGATTTGATGTGACAATGAAATAAGGT

ATGGGAAAGGGAGAACAAACTGATAAATGATAAGTGAATGAAGTAAAGTA

AGGAAGGCTGCCACTATATGCTTTGAGATTTTCATTTAACTATGTTATGCTG

TTGTTAGGGGTTAGAATCTTTTTTCTTGGTTTCCAATTAGGTATGCATTTCTT

TGTATTGGAGAATGTGAGGTTATCTTTTGGGGAGTGACAGCGTTTGGTACAT

TTATTAGGCGATCTATGGTGACATCGCCGCTTGTGCTCCCGTAAGTGCGTTA

GGTGGAGGTGTGCAGGTGTATCAACCTTGAAGGTTGTTCTCGCGCTGCGTTT

TCCATATATAGTAGTTAACTGTGATGCTGCGCATGACTAATTTTTCTGTTTAG

TCACTACACAGTCAGTTAACCTTTTTCCTTTCTGCCACTGTGTCGGGCGCAG

GTTGGTGTTCCACTGGTAACACTCCGTCATAACCAGGCGTGCACGTACCCTC

ACTCCCTGAAATGTGTGACACCGTGGTAAGGCCAACAATATGTCTTATTGAA

ATACTTTCTGGTAATGATGTTGTTGAATATATGTATTGTTTGTCCCCAACCGT

GTTTGTGAATGCTTCATCTCGCTTGTATGCACTCGACTCTCCTCTTTCCACCA

CTCTCCCTAATTTGCTCAACAGTTAAGTCTGTTGTTCATAAAAAACTACGAG

GAAAATACGGAACCGGAGGGGCTCGTGACCGGTGTTGCTCGCTGAGGGAAC

CGTGTGTGAGCGAAGTCTAATATAATATCCGATTCGTCCTTTGGGCGTTTCT

TTCTGCTTATggaggtggatccggaggcggt

Red: End of *Tb*STT3B

Blue: MYC_3 tag

Yellow: Intergenic region

Green: BSD

Purple: 3' UTR

5.2 Acceptor peptide and HA_3 tag for fusion with *Tb*BipN glycosylation reporter protein

ataagtaaCCTAGGGCGGCGGCGACCGAAGGCCTGCTGAACGCGACCGATGAAA

TTGCGCTGCAATTGGCGGCGGCGGAACAGAAACTGATTAGCGAAGAAGATC

TGGAACAGAAACTGATTAGCGAAGAAGATCTGGAACAGAAACTGATTAGC

GAAGAAGATCTGTTAATTAAttacttat

Red: Restriction enzyme sites (AvrII, MfeI and PacI)

Green: Ala_3 linkers

Blue: Acceptor peptide positive control (TEGLLNATDEIAL)

Yellow: HA_3 tag

5.3 Acceptor peptides exclusively acted upon by *Tb*STT3A

*Tb*STT3A acceptor peptides

TAEFRNETAGIAG	SRLEGNDTDCPEE	TILKSNYTAEPVR	PVVVDNRSQAEER
MKASINNTFEGPE	GAAGENSSSPCTV	CTLERNLTVAEML	ELMTRNLSEQESA
MTSVLNTTGYWAS	EIILPNNTEEEEE	EGVKGNLSGADVG	QESAENSTQEKKS
NTWETNYTVFVQQ	EISCTNVTLDEKE	DVGANQTEEEKD	ILAANNYTVNQTD
QTRFDNTTEEKLR	LQEEKNHSEIISE	KAAAANASAALDE	NDKAGNHSAGEDD
SDVAENVTAVKLL	LEDGVNNTDLQDL	TEGLLNATDEIAL	VNAFENTTEGGHH
QLSAANLTDFAHN	TKENFNETIAKSE	VHGNDNVTDQKKG	AETSHNTTLIIEA
PDPPFNFSAVVPE	WEKAANETIDNAL	QVAVRNDTLDFHA	QSARDNATAAAVA
DEMLNNRTDEVVE	PELAANFSIRGYP	DVGALNDTAVLSE	ARLQANHTDCVYT
KALGTNDSYVDPD	GLTANNSTSLST	KAAAANASAALDE	VEIELNGTGLDAT
DTWETNFTVFVQQ	AMVKPNETASIEL	LALFNNETQAVLT	GEVSLNSTFRGRG
VSDETNSSGLYRI	KSNESNDSNESNV	RELISNGSDALDK	AQSYGNESETKTC
PSTPTNSTKNEDS	DKNTLNLTEYFKE	NDLPLNVSREVLQ	IGDEGNETKRTAL
RGIATNETQVEVG	FLHELNSTSSATK	MKEQGNVSEEVNAE	VARNENKTLEQAM
VKRRRNNSQDVDT	DAAVQNKSELPIV	ERDVSNSTVVIAM	KKAEDNVTHLKDY
PATPTNSTKNEDS	VELTDNNSAKCWI	PVYRLNESVCYLK	CEEAENETRDDLR
EKMFRNETEFEKF	KIGVVNGSDMENI	EEVYANASRIIDN	TDVYNNESDGGKY
VVTPENISTEIVA	RQLIGNITVPEQS	LFAEVNESVRAVR	NYEELNVSLSTLE
EEIEENETTGKSE	FMDDPNITEGRME	DEVLENKSTSCLL	WFKNFNYTISGLL
VARNENKTLEEAM	RFAPTNASTTAHT	FLEKQNASDEDIV	VESGANSTRGPSS
SVDHANMSCGSYG	LVRLDNASDPKQV	KPESVNESVNTIL	QRQPANATGEVAP
MCDPENVTYGHTE	EGILENTTALEPY	NDVIPNATTTSRS	NDAMSNSTRAMLD
TAGDYNLTMIEIY	NKNTTNTTSGSDD	YNRLSNATEQLQE	

5.3 Acceptor peptides exclusively acted upon by *Tb*STT3B

*Tb*STT3A acceptor peptides

VSNTFNKTGWDVV	KFLILNLSMKQGL	KLNHTNNSHIIRI
LMRYNNYTKDPFS	KMLTGNLSGFECA	AVVKRNTSLVHVL
LLKLYNFSFVVME	TSSALNRTRQKVG	RSVSSNTSSMHFS
GGSKGNVSGGFVD	KKCKFNATKAEKS	AHLGINVTARTSS
TVRLVNKTGIYTI	AAMIPNATKGTEA	KRNALNATAANKV
VVLWVNISKHPVW	KTGNTNTTGSSNS	PISFKNLTPSELR
NLILRNKTEDALT	KHCKFNSTKAKEK	VKIIENITSAREN
RIGNVNITGVKGA	SYHVVNASLSIGQ	KFKLSNASVGAVR
VSSKSNGTLVSFQ		

6 – References

References

Van Den Abbeele, J., Claes, Y., van Bockstaele, D., Le Ray, D., and Coosemans, M. (1999). Trypanosoma brucei spp. development in the tsetse fly: characterization of the post-mesocyclic stages in the foregut and proboscis. Parasitology *118 (Pt 5*, 469–478).

Acosta-Serrano, A., Rear, J.O., Quellhorst, G., Lee, S.H., Hwa, K., Krag, S.S., Paul, T., and Englund, P.T. (2004). Defects in the N-Linked oligosaccharide biosynthetic pathway in a Trypanosoma brucei glycosylation mutant defects in the N-linked oligosaccharide biosynthetic pathway in a Trypanosoma brucei glycosylation mutant †.

Adung'a, V.O., Gadelha, C., and Field, M.C. (2013). Proteomic analysis of clathrin interactions in trypanosomes reveals dynamic evolution of endocytosis. Traffic *14*, 440–457.

Aebi, M., Gassenhuber, J., Domdey, H., and te Heesen, S. (1996). Cloning and characterization of the ALG3 gene of Saccharomyces cerevisiae. Glycobiology *6*, 439–444.

Apweiler, R., Hermjakob, H., and Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim. Biophys. Acta *1473*, 4–8.

Atrih, A., Richardson, J.M., Prescott, A.R., and Ferguson, M.A.J. (2005). Trypanosoma brucei glycoproteins contain novel giant poly-N-acetyllactosamine carbohydrate chains. J. Biol. Chem. *280*, 865–871.

Balmer, O., Beadell, J.S., Gibson, W., and Caccone, A. (2011). Phylogeography and taxonomy of Trypanosoma brucei. PLoS Negl. Trop. Dis. *5*, e961.

Bandini, G., Mariño, K., Güther, M.L.S., Wernimont, A.K., Kuettel, S., Qiu, W., Afzal, S., Kelner, A., Hui, R., and Ferguson, M.A.J. (2012). Phosphoglucomutase is absent in Trypanosoma brucei and redundantly substituted by phosphomannomutase and phospho-N-acetylglucosamine mutase. Mol. Microbiol. *85*, 513–534.

Bangs, J.D., Doerings, T.L., Englund, P.T., and Hartl, G.W. (1988). Biosynthesis of a Variant Surface Glycoprotein of Trypanosoma brucei. *263*, 17697–17705.

Bangs, J.D., Brouch, E.M., Ransom, D.M., and Roggy, J.L. (1996). A Soluble Secretory Reporter System in Trypanosoma brucei: Studies on endoplasmic reticulum targeting. J. Biol. Chem. *271*, 18387–18393.

Van Berkel, M.A., Rieger, M., te Heesen, S., Ram, A.F., van den Ende, H., Aebi, M., and Klis, F.M. (1999). The Saccharomyces cerevisiae CWH8 gene is required for full levels of dolichol-linked oligosaccharides in the endoplasmic reticulum and for efficient N-glycosylation. Glycobiology *9*, 243–253.

Berriman, M. (2005). The Genome of the African Trypanosome Trypanosoma brucei. Science *309*, 416–422.

Bouteille, B., Oukem, O., Bisser, S., and Dumas, M. (2003). Treatment perspectives for human African trypanosomiasis. Fundam. Clin. Pharmacol. *17*, 171–181.

Broadhead, R., Dawe, H.R., Farr, H., Griffiths, S., Hart, S.R., Portman, N., Shaw, M.K., Ginger, M.L., Gaskell, S.J., McKean, P.G. (2006). Flagellar motility is required for the viability of the bloodstream trypanosome. Nature *440*, 224–227.

Burda, P., and Aebi, M. (1998). The ALG10 locus of Saccharomyces cerevisiae encodes the alpha-1,2 glucosyltransferase of the endoplasmic reticulum: the terminal glucose of the lipid-linked oligosaccharide is required for efficient N-linked glycosylation. Glycobiology *8*, 455–462.

Burda, P., Aebi, M. (1999). The dolichol pathway of N-linked glycosylation. Biochim. Biophys. Acta *1426*, 239–257.

Burda, P., Jakob, C.A., Beinhauer, J., Hegemann, J.H., and Aebi, M. (1999). Ordered assembly of the asymmetrically branched lipid-linked oligosaccharide in the endoplasmic reticulum is ensured by the substrate specificity of the individual glycosyltransferases. Glycobiology *9*, 617–625.

Cáceres, A.J., Michels, P.A.M., and Hannaert, V. (2010). Genetic validation of aldolase and glyceraldehyde-3-phosphate dehydrogenase as drug targets in Trypanosoma brucei. Mol. Biochem. Parasitol. *169*, 50–54.

Castillo-Acosta, V.M., Vidal, A.E., Ruiz-Pérez, L.M., Van Damme, E.J.M., Igarashi, Y., Balzarini, J., and González-Pacanowska, D. (2013). Carbohydrate-binding agents act as potent trypanocidals that elicit modifications in VSG glycosylation and reduced virulence in Trypanosoma brucei. Mol. Microbiol. *90*, 665–679.

Castro, O., Movsichoff, F., and Parodi, A.J. (2006). Preferential transfer of the complete glycan is determined by the oligosaccharyltransferase complex and not by the catalytic subunit. Proc. Natl. Acad. Sci. U. S. A. *103*, 14756–14760.

Chaban, B., Voisin, S., Kelly, J., Logan, S.M., and Jarrell, K.F. (2006). Identification of genes involved in the biosynthesis and attachment of Methanococcus voltae N-linked glycans: insight into N-linked glycosylation pathways in Archaea. Mol. Microbiol. *61*, 259–268.

Chappuis, F., Loutan, L., Simarro, P., Büscher, P., Lejon, V., and Bu, P. (2005a). Options for Field Diagnosis of Human African Trypanosomiasis Options for Field Diagnosis of Human African Trypanosomiasis Franc. *18*.

Chappuis, F., Udayraj, N., Stietenroth, K., Meussen, A., and Bovier, P.A. (2005b). Eflornithine is safer than melarsoprol for the treatment of second-stage Trypanosoma brucei gambiense human African trypanosomiasis. Clin. Infect. Dis. *41*, 748–751.

Checchi, F., Filipe, J.A.N., Haydon, D.T., Chandramohan, D., and Chappuis, F. (2008a). Estimates of the duration of the early and late stage of gambiense sleeping sickness. BMC Infect. Dis. *8*, 16.

Checchi, F., Filipe, J.A.N., Barrett, M.P., and Chandramohan, D. (2008b). The natural progression of Gambiense sleeping sickness: what is the evidence? PLoS Negl. Trop. Dis. *2*, e303.

Chen, M.M., Glover, K.J., and Imperiali, B. (2007). From Peptide to Protein : Comparative Analysis of the Substrate Specificity of. Biochemistry *46*, 5579–5585.

Cherepanova, N.A., Shrimal, S., and Gilmore, R. (2014). Oxidoreductase activity is necessary for N-glycosylation of cysteine-proximal acceptor sites in glycoproteins. J. Cell Biol. *206*, 525–539.

Choi, K.-J., Grass, S., Paek, S., St Geme, J.W., and Yeo, H.-J. (2010). The Actinobacillus pleuropneumoniae HMW1C-like glycosyltransferase mediates N-linked glycosylation of the Haemophilus influenzae HMW1 adhesin. PLoS One *5*, e15888.

Cipollo, J.F., Trimble, R.B., Chi, J.H., Yan, Q., Dean, N. (2001). The yeast ALG11 gene specifies addition of the terminal alpha 1,2-Man to the Man5GlcNAc2-PP-dolichol N-glycosylation intermediate formed on the cytosolic side of the endoplasmic reticulum. J. Biol. Chem. *276*, 21828–21840.

Couto, J.R., Huffaker, T.C., and Robbins, P.W. (1984). Cloning and expression in Escherichia coli of a yeast mannosyltransferase from the Cloning and Expression in Escherichia coli of a Yeast Mannosyltransferase from the Asparagine-linked Glycosylation Pathway. J. Biol. Chem. *259*, 378–382.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Crimaudo, C., Hortsch, M., Gausepohl, H., and Meyer, D.I. (1987). Human ribophorins I and II: the primary structure and membrane topology of two highly conserved rough endoplasmic reticulum-specific glycoproteins. EMBO J. *6*, 75–82.

Cross, G.A (1975a). Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of Trypanosoma brucei. Parasitology *71*, 393–417.

Cross, G.A. (1975b). Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of Trypanosoma brucei. Parasitology *71*, 393–417.

Cross, G.A. (1979). Crossreacting determinants in the C-terminal region of trypanosome variant surface antigens. Nature *277*, 310–312.

Cross, G.A. (1984). Release and purification of Trypanosoma brucei variant surface glycoprotein. J. Cell. Biochem. *24*, 79–90.

Damerow, M., Rodrigues, J.A., Wu, D., Güther, M.L.S., Mehlert, A., and Ferguson, M.A.J. (2014). Identification and functional characterization of a highly divergent N-acetylglucosaminyltransferase I (TbGnTI) in Trypanosoma brucei. J. Biol. Chem. *289*, 9328–9339.

Dell, A., Galadari, A., Sastre, F., and Hitchen, P. (2010). Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. Int. J. Microbiol. *2010*, 148178.

Dumax-Vorzet, A., Roboti, P., and High, S. (2013). OST4 is a subunit of the mammalian oligosaccharyltransferase required for efficient N-glycosylation. J. Cell Sci. *126*, 2595–2606.

Engstler, M., Weise, F., Bopp, K., Grünfelder, C.G., Günzel, M., Heddergott, N., and Overath, P. (2005). The membrane-bound histidine acid phosphatase TbMBAP1 is essential for endocytosis and membrane recycling in Trypanosoma brucei. J. Cell Sci. *118*, 2105–2118.

Fairlamb, A.H. (2003). Chemotherapy of human African trypanosomiasis: current and future prospects. Trends Parasitol. *19*, 488–494.

Fenn, K., and Matthews, K.R. (2007). The cell biology of Trypanosoma brucei differentiation. Curr. Opin. Microbiol. *10*, 539–546.

Ferguson, M.A.J., Haldar, K., and Cross, G.A. (1985). Trypanosoma brucei variant surface glycoprotein has a sn-1,2-dimyristyl glycerol membrane anchor at its COOH terminus. J. Biol. Chem. *260*, 4963–4968.

Ferguson, M.A.J., Homans, S.W., Dwek, R.A., and Rademacher, T.W. (1988). Glycosyl-phosphatidylinositol moiety that anchors Trypanosoma brucei variant surface glycoprotein to the membrane. Science *239*, 753–759.

Fetrow, J.S., Siew, N., Gennaro, J.A.D.I., Martinez-yamout, M., Dyson, H.J., and Skolnick, J. (2001). Genomic-scale comparison of sequence- and structure-based methods of function prediction□: Does structure provide additional insight□? 1005–1014.

Frank, C.G., and Aebi, M. (2005). ALG9 mannosyltransferase is involved in two different steps of lipid-linked oligosaccharide biosynthesis. Glycobiology *15*, 1156–1163.

Frank, C.G., Sanyal, S., Rush, J.S., Waechter, C.J., and Menon, A.K. (2008). Does Rft1 flip an N-glycan lipid precursor? Nature *454*, E3–E4; discussion E4–E5.

Fu, J., Ren, M., and Kreibich, G. (1997). Interactions among Subunits of the Oligosaccharyltransferase Complex. J. Biol. Chem. *272*, 29687–29692.

Gayen, S., and Kang, C. (2011). Solution structure of a human minimembrane protein Ost4, a subunit of the oligosaccharyltransferase complex. Biochem. Biophys. Res. Commun. *409*, 572–576.

Grass, S., Buscher, A.Z., Swords, W.E., Apicella, M.A, Barenkamp, S.J., Ozchlewski, N., and St Geme, J.W. (2003). The Haemophilus influenzae HMW1 adhesin is glycosylated in a process that requires HMW1C and phosphoglucomutase, an enzyme involved in lipooligosaccharide biosynthesis. Mol. Microbiol. *48*, 737–751.

Grass, S., Lichti, C.F., Townsend, R.R., Gross, J., and St Geme, J.W. (2010). The Haemophilus influenzae HMW1C protein is a glycosyltransferase that transfers hexose residues to asparagine sites in the HMW1 adhesin. PLoS Pathog. *6*, e1000919.

Gross, J., Grass, S., Davis, A.E., Gilmore-Erdmann, P., Townsend, R.R., and St Geme, J.W. (2008). The Haemophilus influenzae HMW1 adhesin is a glycoprotein with an unusual N-linked carbohydrate modification. J. Biol. Chem. *283*, 26010–26015.

Guan, Z., Naparstek, S., Kaminski, L., Konrad, Z., and Eichler, J. (2010). Distinct glycan-charged phosphodolichol carriers are required for the assembly of the pentasaccharide N-linked to the Haloferax volcanii S-layer glycoprotein. Mol. Microbiol. *78*, 1294–1303.

Güther, M.L.S., Prescott, A.R., and Ferguson, M.A.J. (2003). Deletion of the GPIdeAc gene alters the location and fate of glycosylphosphatidylinositol precursors in Trypanosoma brucei. Biochemistry *42*, 14532–14540.

Heesen, S., Lehle, L., Weissmann, A, and Aebi, M. (1994). Isolation of the ALG5 locus encoding the UDP-glucose:dolichyl-phosphate glucosyltransferase from Saccharomyces cerevisiae. Eur. J. Biochem. *224*, 71–79.

Helenius, A. (1994). Essay How N-linked Oligosaccharides Affect Glycoprotein Folding in the Endoplasmic Reticulum. *5*, 253–265.

Helenius, J., and Aebi, M. (2002). Transmembrane movement of dolichol linked carbohydrates during N -glycoprotein biosynthesis in the endoplasmic reticulum. *13*, 171–178.

Heller, L., Orlean, P., and Adair, W.L. (1992). Saccharomyces cerevisiae sec59 cells are deficient in dolichol kinase activity. Proc. Natl. Acad. Sci. U. S. A. *89*, 7013–7016.

Hese, K., Otto, C., Routier, F.H., and Lehle, L. (2009). The yeast oligosaccharyltransferase complex can be replaced by STT3 from Leishmania major. Glycobiology *19*, 160–171.

Hirumi H, Hirumi, K. (1989). Continuous Cultivation of Trypanosoma brucei Blood Stream Forms in a Medium Containing a Low Concentration of Serum Protein without Feeder Cell Layers Hiroyuki Hirumi□; Kazuko Hirumi Continuous Cultivation of Trypanosomabrucei Blood Stream Forms in a Medi. J. Parasitol. *75*, 985–989.

Honma, K., Iwao-Koizumi, K., Takeshita, F., Yamamoto, Y., Yoshida, T., Nishio, K., Nagahara, S., Kato, K., and Ochiya, T. (2008). RPN2 gene confers docetaxel resistance in breast cancer. Nat. Med. *14*, 939–948.

Horn, D., and Cross, G.A. (1997). Analysis of Trypanosoma brucei vsg expression site switching in vitro. Mol. Biochem. Parasitol. *84*, 189–201.

Huffaker, T.C., and Robbins, P.W. (1983). Yeast mutants deficient in protein glycosylation. Proc. Natl. Acad. Sci. U. S. A. *80*, 7466–7470.

Izquierdo, L., Schulz, B.L., Rodrigues, J.A., Güther, M.L.S., Procter, J.B., Barton, G.J., Aebi, M., and Ferguson, M.A.J. (2009a). Distinct donor and acceptor specificities of Trypanosoma brucei oligosaccharyltransferases. EMBO J. *28*, 2650–2661.

Izquierdo, L., Atrih, A., Rodrigues, J. a, Jones, D.C., and Ferguson, M. a J. (2009b). Trypanosoma brucei UDP-glucose:glycoprotein glucosyltransferase has unusual substrate specificity and protects the parasite from stress. Eukaryot. Cell *8*, 230–240.

Izquierdo, L., Mehlert, A., and Ferguson, M.A.J. (2012). The lipid-linked oligosaccharide donor specificities of Trypanosoma brucei oligosaccharyltransferases. Glycobiology *22*, 696–703.

Jelk, J., Gao, N., Serricchio, M., Signorell, A., Schmidt, R.S., Bangs, J.D., Acosta-Serrano, A., Lehrman, M.A., Bütikofer, P., and Menon, A.K. (2013). Glycoprotein biosynthesis in a eukaryote lacking the membrane protein Rft1. J. Biol. Chem. *288*, 20616–20623.

Jenkins, N., Parekh, R.B., and James, D.C. (1996). Getting the glycosylation right: Implications for the biotechnology industry. Nat. Biotechnol. *14*, 975–981.

Jervis, A.J., Langdon, R., Hitchen, P., Lawson, A.J., Wood, A., Fothergill, J.L., Morris, H.R., Dell, A., Wren, B., and Linton, D. (2010). Characterization of N-linked protein glycosylation in Helicobacter pullorum. J. Bacteriol. *192*, 5228–5236.

Jones, D.C., Mehlert, A., Güther, M.L.S., and Ferguson, M. a J. (2005). Deletion of the glucosidase II gene in Trypanosoma brucei reveals novel N-glycosylation mechanisms in the biosynthesis of variant surface glycoprotein. J. Biol. Chem. *280*, 35929–35942.

Kelleher, D.J., and Gilmore, R. (1997). DAD1, the defender against apoptotic cell death, is a subunit of the mammalian oligosaccharyltransferase. Proc. Natl. Acad. Sci. U. S. A. *94*, 4994–4999.

Kelleher, D.J., and Gilmore, R. (2006). An evolving view of the eukaryotic oligosaccharyltransferase. Glycobiology *16*, 47R – 62R.

Kelleher, D.J., Karaoglu, D., Mandon, E.C., and Gilmore, R. (2003). Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. Mol. Cell *12*, 101–111.

Kelleher, D.J., Banerjee, S., Cura, A.J., Samuelson, J., and Gilmore, R. (2007). Dolichol-linked oligosaccharide selection by the oligosaccharyltransferase in protist and fungal organisms. J. Cell Biol. *177*, 29–37.

Kelley, L.A., and Sternberg, M.J.E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protoc. *4*, 363–371.

Kelley, R.J., Alexander, D.L., Cowan, C., Balber, A.E., and Bangs, J.D. (1999). Molecular cloning of p67, a lysosomal membrane glycoprotein from Trypanosoma brucei. Mol. Biochem. Parasitol. *98*, 17–28.

Kelly, J., Logan, S.M., Jarrell, K.F., VanDyke, D.J., and Vinogradov, E. (2009). A novel N-linked flagellar glycan from Methanococcus maripaludis. Carbohydr. Res. *344*, 648–653.

Kennedy, P.G.E. (2012). An alternative form of melarsoprol in sleeping sickness. Trends Parasitol. *28*, 307–310.

Khoury, G. a, Baliban, R.C., and Floudas, C. a (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. Sci. Rep. *1*, 1–5.

Kim, H., Park, H., Montalvo, L., and Lennarz, W.J. (2000). Studies on the role of the hydrophobic domain of Ost4p in interactions with other subunits of yeast oligosaccharyl transferase. Proc. Natl. Acad. Sci. U. S. A. *97*, 1516–1520.

Kim, H., Yan, Q., Heijne, G. Von, Caputo, G.A., and Lennarz, W.J. (2003). Determination of the membrane topology of Ost4p and its subunit interactions in the oligosaccharyl- transferase complex in Saccharomyces cerevisiae.

Kim, H., von Heijne, G., and Nilsson, I. (2005). Membrane topology of the STT3 subunit of the oligosaccharyl transferase complex. J. Biol. Chem. *280*, 20261–20267.

Kowarik, M., Küng, S., Martoglio, B., and Helenius, A. (2002). Protein folding during cotranslational translocation in the endoplasmic reticulum. Mol. Cell *10*, 769–778.

Kreibich, G., Ulrich, B.L., and Sabanti, D.D. (1978a). Proteinsofroughmembranesrelatedtoribosomebindning I . Identificationof Ribophorins I and II , Membrane Proteinscharacteristic of Rough Microsomes Rat liver rough microsomes (RM) contain two in. J. CELL Biol. Rockefeller Univ. Press *77*, 464–487.

Kreibich, G., Freienstein, C.M., Pereyra, B.N., Ulrich, B.L., and Sabanti, D.D. (1978b). Proteinsofroughmembranesrelatedtoribosomebindning II . Cross-linking ofbound ribosomes to specific membraneproteins. J. CELL Biol. Rockefeller Univ. Press *77*, 488–506.

Kuettel, S., Wadum, M.C.T., Güther, M.L.S., Mariño, K., Riemer, C., and Ferguson, M.A.J. (2012). The de novo and salvage pathways of GDP-mannose biosynthesis are both sufficient for the growth of bloodstream-form Trypanosoma brucei. Mol. Microbiol. *84*, 340–351.

Landfear, S.M., and Ignatushchenko, M. (2001). The flagellum and flagellar pocket of trypanosomatids. Mol. Biochem. Parasitol. *115*, 1–17.

Lehrman, M.A. (1991). Biosynthesis of N-acetylglucosamine-P-P-dolichol, the committed step of asparagine-linked oligosaccharide assembly. Glycobiology *1*, 553–562.

Lejon, V., Ngoyi, D.M., Boelaert, M., and Büscher, P. (2010). A CATT negative result after treatment for human African trypanosomiasis is no indication for cure. PLoS Negl. Trop. Dis. *4*, e590.

Li, H., Chavan, M., Schindelin, H., Lennarz, W.J., and Li, H. (2008). Structure of the oligosaccharyl transferase complex at 12 A resolution. Structure *16*, 432–440.

Lingnau, a, Zufferey, R., Lingnau, M., and Russell, D.G. (1999). Characterization of tGLP-1, a Golgi and lysosome-associated, transmembrane glycoprotein of African trypanosomes. J. Cell Sci. *112 Pt 18*, 3061–3070.

Lizak, C., Gerber, S., Numao, S., Aebi, M., and Locher, K.P. (2011). X-ray structure of a bacterial oligosaccharyltransferase. Nature *474*, 350–355.

Lizak, C., Gerber, S., Zinne, D., Michaud, G., Schubert, M., Chen, F., Bucher, M., Darbre, T., Zenobi, R., Reymond, J.-L. (2014). A catalytically essential motif in external loop 5 of the bacterial oligosaccharyltransferase PglB. J. Biol. Chem. *289*, 735–746.

Low, P., Dallnersy, G., Mayorii, S., Cohen, S., Chait, B.T., and Menon, A.K. (1991). The Mevalonate Pathway in the Bloodstream Form of Trypanosoma brucei. 19250–19257.

MacGrogan, D., Levy, A., Bova, G.S., Isaacs, W.B., and Bookstein, R. (1996). Structure and methylation-associated silencing of a gene within a homozygously deleted region of human chromosome band 8p22. Genomics *35*, 55–65.

MacLean, L., Reiber, H., Kennedy, P.G.E., and Sternberg, J.M. (2012). Stage progression and neurological symptoms in Trypanosoma brucei rhodesiense sleeping sickness: role of the CNS inflammatory response. PLoS Negl. Trop. Dis. *6*, e1857.

Magidovich, H., and Eichler, J. (2009). Glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of the N-glycosylation pathway in the third domain of life. FEMS Microbiol. Lett. *300*, 122–130.

Malaby, H.L.H., and Kobertz, W.R. (2014). The middle x residue influences cotranslational N-glycosylation consensus site skipping. Biochemistry *53*, 4884–4893.

Maley, F., Trimble, R.B., Tarentino, A.L., and Plummer, T.H. (1989). Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. Anal. Biochem. *180*, 195–204.

Manthri, S., Güther, M.L.S., Izquierdo, L., Acosta-Serrano, A., and Ferguson, M.A.J. (2008). Deletion of the TbALG3 gene demonstrates site-specific N-glycosylation and N-glycan processing in Trypanosoma brucei. Glycobiology *18*, 367–383.

Marcantonio, E.E., Amar-Costesec, a, and Kreibich, G. (1984). Segregation of the polypeptide translocation apparatus to regions of the endoplasmic reticulum containing ribophorins and ribosomes. II. Rat liver microsomal subfractions contain equimolar amounts of ribophorins and ribosomes. J. Cell Biol. *99*, 2254–2259.

Mariño, K., Güther, M.L.S., Wernimont, A.K., Amani, M., Hui, R., and Ferguson, M.A.J. (2010). Identification, subcellular localization, biochemical properties, and high-resolution crystal structure of Trypanosoma brucei UDP-glucose pyrophosphorylase. Glycobiology *20*, 1619–1630.

Mehlert, A., Zitzmann, N., Richardson, J.M., Truemann, A., and Ferguson, M.A.J. (1998). The glycosylation of the variant surface glycoproteins and procyclic acidic repetitive proteins of Trypanosoma brucei. Mol. Biochem. Parasitol. *91*, 145–152.

Mehlert, A., Bond, C.S., and Ferguson, M.A.J. (2002). The glycoforms of a Trypanosoma brucei variant surface glycoprotein and molecular modeling of a glycosylated surface coat. Glycobiology *12*, 607–612.

Mehlert, A., Sullivan, L., and Ferguson, M.A.J. (2010). Glycotyping of Trypanosoma brucei variant surface glycoprotein MITat1.8. Mol. Biochem. Parasitol. *174*, 74–77.

Mehlert, A., Wormald, M.R., and Ferguson, M.A.J. (2012). Modeling of the N-glycosylated transferrin receptor suggests how transferrin binding can occur within the surface coat of Trypanosoma brucei. PLoS Pathog. *8*, e1002618.

Michels, P.A.M., Bringaud, F., Herman, M., and Hannaert, V. (2006). Metabolic functions of glycosomes in trypanosomatids. Biochim. Biophys. Acta *1763*, 1463–1477.

Mohorko, E., Glockshuber, R., and Aebi, M. (2011). Oligosaccharyltransferase: the central enzyme of N-linked protein glycosylation. J. Inherit. Metab. Dis. *34*, 869–878.

Mohorko, E., Owen, R.L., Malojčić, G., Brozzo, M.S., Aebi, M., and Glockshuber, R. (2014). Structural basis of substrate specificity of human oligosaccharyl transferase subunit N33/Tusc3 and its role in regulating protein N-glycosylation. Structure *22*, 590–601.

Nagamune, K., Ohishi, K., Ashida, H., Hong, Y., Hino, J., Kangawa, K., Inoue, N., Maeda, Y., and Kinoshita, T. (2003). GPI transamidase of Trypanosoma brucei has two previously uncharacterized (trypanosomatid transamidase 1 and 2) and three common subunits. Proc. Natl. Acad. Sci. U. S. A. *100*, 10682–10687.

Nakanishi, M., Karasudani, M., Shiraishi, T., Hashida, K., Hino, M., Ferguson, M.A.J., and Nomoto, H. (2014). TbGT8 is a bifunctional glycosyltransferase that elaborates N-linked glycans on a protein phosphatase AcP115 and a GPI-anchor modifying glycan in Trypanosoma brucei. Parasitol. Int. *63*, 513–518.

Nasab, F.P., Schulz, B.L., Gamarro, F., Parodi, A.J., and Aebi, M. (2008). All in One□: Leishmania major STT3 Proteins Substitute for the Whole Oligosaccharyltransferase Complex in Saccharomyces cerevisiae. *19*, 3758–3768.

Ng, D.T., Spear, E.D., and Walter, P. (2000). The unfolded protein response regulates multiple aspects of secretory and membrane protein biogenesis and endoplasmic reticulum quality control. J. Cell Biol. *150*, 77–88.

Nilsson, I., Kelleher, D.J., Miao, Y., Shao, Y., Kreibich, G., Gilmore, R., von Heijne, G., and Johnson, A.E. (2003). Photocross-linking of nascent chains to the STT3 subunit of the oligosaccharyltransferase complex. J. Cell Biol. *161*, 715–725.

Nozaki, T. (1996). Characterization of the Trypanosoma brucei homologue of a Trypanosoma cruzi flagellum-adhesion glycoprotein. Mol. Biochem. Parasitol. *82*, 245–255.

O'Reilly, M.K.O., Zhang, G., and Imperiali, B. (2006). In Vitro Evidence for the Dual Function of Alg2 and Alg11□: Essential. *14*, 9593–9603.

Oberholzer, M., Morand, S., Kunz, S., and Seebeck, T. (2006). A vector series for rapid PCR-mediated C-terminal in situ tagging of Trypanosoma brucei genes. Mol. Biochem. Parasitol. *145*, 117–120.

Oeljeklaus, S., Reinartz, B.S., Wolf, J., Wiese, S., Tonillo, J., Podwojski, K., Kuhlmann, K., Stephan, C., Meyer, H.E., Schliebs, W., et al. (2012). Identification of core components and transient interactors of the peroxisomal importomer by dual-track stable isotope labeling with amino acids in cell culture analysis. J. Proteome Res. *11*, 2567–2580.

Opperdoes, F.R., and Borst, P. (1977). Localization of nine glycolytic microbody-like organelle in Trypanosoma brucei: the glycosome. FEBS lett.*80,* 360-4.

Orlean, P., Albright, C., and Robbins, P.W. (1988). Cloning and sequencing of the yeast gene for dolichol phosphate mannose synthase, an essential protein. J. Biol. Chem. *263*, 17499–17507.

Paquet, C., Ancelle, T., Gastellu-Etchegorry, M., Castilla, J., and Harndt, I. (1992). Persistence of antibodies to Trypanosoma brucei gambiense after treatment of human trypanosomiasis in Uganda. Lancet *340*, 250.

Pathak, R., Hendrickson, T.L., and Imperiali, B. (1995). Sulfhydryl modification of the yeast Wbp1p inhibits oligosaccharyl transferase activity. Biochemistry *34*, 4179–4185.

Petrescu, A.-J., Milac, A.-L., Petrescu, S.M., Dwek, R.A., and Wormald, M.R. (2004). Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. Glycobiology *14*, 103–114.

Priotto, G., Kasparian, S., Mutombo, W., Ngouama, D., Ghorashian, S., Arnold, U., Ghabri, S., Baudin, E., Buard, V., Kazadi-Kyanza, S. (2009). Nifurtimox-eflornithine combination therapy for second-stage African Trypanosoma brucei gambiense trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. Lancet *374*, 56–64.

Reiss, G., Heesen, S., Zimmerman, J., Robbins, P.W., and Aebi, M. (1996). Isolation of the ALG6 locus of Saccharomyces cerevisiae required for glucosylation in the N-linked glycosylation pathway. Glycobiology *6*, 493–498.

Renslo, A.R., and McKerrow, J.H. (2006). Drug discovery and development for neglected parasitic diseases. Nat. Chem. Biol. *2*, 701–710.

Roboti, P., and High, S. (2012a). The oligosaccharyltransferase subunits OST48, DAD1 and KCP2 function as ubiquitous and selective modulators of mammalian N-glycosylation. J. Cell Sci. *125*, 3474–3484.

Roper, J.R., Güther, M.L.S., Macrae, J.I., Prescott, A.R., Hallyburton, I., Acosta-Serrano, A., and Ferguson, M.A.J. (2005). The suppression of galactose metabolism in procylic form Trypanosoma brucei causes cessation of cell growth and alters procyclin glycoprotein structure and copy number. J. Biol. Chem. *280*, 19728–19736.

Ruiz-Canada, C., Kelleher, D.J., and Gilmore, R. (2009). Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. Cell *136*, 272–283.

Rush, J.S., Gao, N., Lehrman, M.A., Matveev, S., and Waechter, C.J. (2009). Suppression of Rft1 expression does not impair the transbilayer movement of Man5GlcNAc2-P-P-dolichol in sealed microsomes from yeast. J. Biol. Chem. *284*, 19835–19842.

Samuelson, J., Banerjee, S., Magnelli, P., Cui, J., Kelleher, D.J., Gilmore, R., and Robbins, P.W. (2005). The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases. Proc. Natl. Acad. Sci. U. S. A. *102*, 1548–1553.

Sbicego, S., Vassella, E., Kurath, U., Blum, B., and Roditi, I. (1999). The use of transgenic Trypanosoma brucei to identify compounds inducing the differentiation of bloodstream forms to procyclic forms. Mol. Biochem. Parasitol. *104*, 311–322.

Schenk, B., Fernandez, F., and Waechter, C.J. (2001). The ins (ide) and outs (ide) of dolichyl phosphate biosynthesis and recycling in the endoplasmic reticulum The precursor oligosaccharide donor for protein N-glyco-. *11*.

Schulz, B.L., Stirnimann, C.U., Grimshaw, J.P.A., Brozzo, M.S., Fritsch, F., Mohorko, E., Capitani, G., Glockshuber, R., Grütter, M.G., and Aebi, M. (2009). Oxidoreductase activity of oligosaccharyltransferase subunits Ost3p and Ost6p defines site-specific glycosylation efficiency. Proc. Natl. Acad. Sci. U. S. A. *106*, 11061–11066.

Schwarz, F., Lizak, C., Fan, Y.-Y., Fleurkens, S., Kowarik, M., and Aebi, M. (2011). Relaxed acceptor site specificity of bacterial oligosaccharyltransferase in vivo. Glycobiology *21*, 45–54.

Sharma, R., Gluenz, E., Peacock, L., Gibson, W., Gull, K., and Carrington, M. (2009). The heart of darkness: growth and form of Trypanosoma brucei in the tsetse fly. Trends Parasitol. *25*, 517–524.

Sharon, N., and Lis, H. (1995). Lectins--proteins with a sweet tooth: functions in cell recognition. Essays Biochem. *30*, 59–75.

Shibatani, T., David, L.L., McCormack, A.L., Frueh, K., and Skach, W.R. (2005). Proteomic analysis of mammalian oligosaccharyltransferase reveals multiple subcomplexes that contain Sec61, TRAP, and two potential new subunits. Biochemistry *44*, 5982–5992.

Silberstein, S., Kelleher, D.J., and Gilmore, R. (1992). The 48-kDa subunit of the mammalian oligosaccharyltransferase complex is homologous to the essential yeast protein WBP1. J. Biol. Chem. *267*, 23658–23663.

Simarro, P.P., Jannin, J., and Cattand, P. (2008). Eliminating Human African Trypanosomiasis□: Where Do We Stand and What Comes Next□? *5*, 174–180.

Simarro, P.P., Cecchi, G., Paone, M., Franco, J.R., Diarra, A., Ruiz, J.A., Fèvre, E.M., Courtin, F., Mattioli, R.C., and Jannin, J.G. (2010). The Atlas of human African

trypanosomiasis: a contribution to global mapping of neglected tropical diseases. Int. J. Health Geogr. *9*, 57.

Simarro, P.P., Cecchi, G., Franco, J.R., Paone, M., Fèvre, E.M., Diarra, A., Antonio, J., Postigo, R., Mattioli, R.C., and Jannin, J.G. (2011). Risk for Human African. *17*, 5–7.

Simarro, P.P., Cecchi, G., Franco, J.R., Paone, M., Diarra, A., Ruiz-Postigo, J.A., Fèvre, E.M., Mattioli, R.C., and Jannin, J.G. (2012). Estimating and mapping the population at risk of sleeping sickness. PLoS Negl. Trop. Dis. *6*, e1859.

Singha, U.K., Hamilton, V., Duncan, M.R., Weems, E., Tripathi, M.K., and Chaudhuri, M. (2012). Protein translocase of mitochondrial inner membrane in Trypanosoma brucei. J. Biol. Chem. *287*, 14480–14493.

Spirig, U., Glavas, M., Bodmer, D., Reiss, G., Burda, P., Lippuner, V., te Heesen, S., and Aebi, M. (1997). The STT3 protein is a component of the yeast oligosaccharyltransferase complex. Mol. Gen. Genet. *256*, 628–637.

Stagljar, I., te Heesen, S., and Aebi, M. (1994). New phenotype of mutations deficient in glucosylation of the lipid-linked oligosaccharide: cloning of the ALG8 locus. Proc. Natl. Acad. Sci. U. S. A. *91*, 5977–5981.

Steverding, D., Stierhof, Y.D., Chaudhri, M., Ligtenberg, M., Schell, D., Beck-Sickinger, A.G., and Overath, P. (1994). ESAG 6 and 7 products of Trypanosoma brucei form a transferrin binding protein complex. Eur. J. Cell Biol. *64*, 78–87.

Stokes, M.J., Güther, M.L.S., Turnock, D.C., Prescott, A.R., Martin, K.L., Alphey, M.S., and Ferguson, M.A.J. (2008). The synthesis of UDP-N-acetylglucosamine is essential for bloodstream form trypanosoma brucei in vitro and in vivo and UDP-N-acetylglucosamine starvation reveals a hierarchy in parasite protein glycosylation. J. Biol. Chem. *283*, 16147–16161.

Strang, A.M., Allen, A.K., Holder, A.A., and Vanhalbeek, H. (1993). The Carbohydrate Structures of Trypanosoma brucei brucei MITat 1.6 Variant Surface Glycoprotein: A Reinvestigation of the C-Terminal Glycan. Biochem. Biophys. Res. Commun. *196*, 1430–1439.

Swiezewska, E., and Danikiewicz, W. (2005). Polyisoprenoids: structure, biosynthesis and function. Prog. Lipid Res. *44*, 235–258.

Turnock, D.C., and Ferguson, M.A.J. (2007). Sugar nucleotide pools of Trypanosoma brucei, Trypanosoma cruzi, and Leishmania major. Eukaryot. Cell *6*, 1450–1463.

Urbaniak, M.D., Martin, D.M. a, and Ferguson, M.A.J. (2013). Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of Trypanosoma brucei. J. Proteome Res. *12*, 2233–2244.

Varki, A., Cummings, R., E., Freeze, J., H., Hart, G., A., and Marth, J. (1999). Essentials of glycobiology.

Varki A, Esko JD, C.K. (2009). Cellular Organization of Glycosylation - Essentials of Glycobiology. In Essentials of Glycobiology. 2nd Edition. Chapter 3.

Wacker, M., Feldman, M.F., Callewaert, N., Kowarik, M., Clarke, B.R., Pohl, N.L., Hernandez, M., Vines, E.D., Valvano, M.A., Whitfield, C. (2006). Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. Proc. Natl. Acad. Sci. U. S. A. *103*, 7088–7093.

Welti, M. (2013). Regulation of dolichol-linked glycosylation. Glycoconj. J. *30*, 51–56.

Wilson, C.M., and High, S. (2007). Ribophorin I acts as a substrate-specific facilitator of N-glycosylation. J. Cell Sci. *120*, 648–657.

Wilson, C.M., Kraft, C., Duggan, C., Ismail, N., Crawshaw, S.G., and High, S. (2005). Ribophorin I associates with a subset of membrane proteins after their integration at the sec61 translocon. J. Biol. Chem. *280*, 4195–4206.

Wirtz, E., Leal, S., Ochatt, C., and Cross, G.A (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Mol. Biochem. Parasitol. *99*, 89–101.

Wolburg, H., Mogk, S., Acker, S., Frey, C., Meinert, M., Schönfeld, C., Lazarus, M., Urade, Y., Kubata, B.K., and Duszenko, M. (2012). Late stage infection in sleeping sickness. PLoS One *7*, e34304.

Wooten, E.W., Bazzo, R., Edge, C.J., Zamze, S., Dwek, R.A., and Rademacher, T.W. (1990). Primary sequence dependence of conformation in oligomannose oligosaccharides. Eur. Biophys. J. *18*, 139–148.

Wormald, M.R., and Dwek, R.A. (1999). Glycoproteins: glycan presentation and protein-fold stability. Structure *7*, R155–R160.

Woycechowsky, K.J., and Raines, R.T. (2003). The CXC motif: a functional mimic of protein disulfide isomerase. Biochemistry *42*, 5387–5394.

Zamze, S.E., Ashford, D. a, Wooten, E.W., Rademacher, T.W., and Dwek, R.A. (1991). Structural characterization of the asparagine-linked oligosaccharides from Trypanosoma brucei type II and type III variant surface glycoproteins. J. Biol. Chem. *266*, 20244–20261.

Zeitler, R., Hochmuth, E., Deutzmann, R., and Sumper, M. (1998). Exchange of Ser-4 for Val, Leu or Asn in the sequon Asn-Ala-Ser does not prevent N-glycosylation of the cell surface glycoprotein from Halobacterium halobium. Glycobiology *8*, 1157–1164.

Ziegelbauer, K., and Overath, P. (1992). Identification of invariant surface glycoproteins in the bloodstream stage of Trypanosoma brucei. J. Biol. Chem. *267*, 10791–10796.

Zielinska, D.F., Gnad, F., Wiśniewski, J.R., and Mann, M. (2010). Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. Cell *141*, 897–907.

Zielinska, D.F., Gnad, F., Schropp, K., Wiśniewski, J.R., and Mann, M. (2012). Mapping N-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. Mol. Cell *46*, 542–548.

Zomer, A.W., Opperdoes, F.R., and van den Bosch, H. (1995). Alkyl dihydroxyacetone phosphate synthase in glycosomes of Trypanosoma brucei. Biochim. Biophys. Acta *1257*, 167–173.